Misjudgment Explanation in Ambiguous and Comparative Contexts

Gueorgui Petkov[a,b*]
[a]Sofia University St. Kliment Ohridski, Sofia, Bulgaria
[b]Kozloduy Nuclear Power Plant Plc., Kozloduy, Bulgaria

Abstract: Making more rational risk-informed decisions despite human biases may be one of the most important applications of the symptom-based contextualization for probabilistic safety assessment and management. Misjudgment identification is the detection of misrecognition of the quantity or quality of one or more sets of symptoms. This consists in anticipating upcoming variability in symptom recognition, i.e. detecting, complementing, or ignoring additional information that may correct for unexpected ambiguous context. Avoiding cognitive bias-based misjudgments is primarily concerned with limiting the effects of delayed or violated symptom recognition through careful assessment of grouping, qualification, and quantification of symptoms. The paper presents the capabilities of the symptom-based context quantification procedure by the Performance Evaluation of Teamwork method for modeling misjudgment from logical cognitive biases. This method is applicable to improve the human reliability assessment and is useful for understanding the iterativity, repeatability and complementarity of thought processes for their rational modeling and explanation in ambiguous and comparative contexts. The basic idea of overcoming the uncertainty of judgment is to heuristically model the interference of symptom recognition waves and to justify the fact that at any moment the judgment is based on tracking all trajectories of the real context. Each context trajectory is identified by unique combination of partially recognized system's states versus all possible system's states, where the contexture is a ratio of these counted numbers of states by the PET context quantification procedure. The paper shows how the method can be applied to identify and quantify the probability of conjunction and disjunction fallacies or combination thereof. The goal is to implement the methodology for data mining, consisting of probabilistic heuristic modeling, simulating, identifying, resolving and explaining logical errors of commission in the regular training of operator crews.

Keywords: symptom-based context, HRA, risk-informed judgment, errors under uncertainty, logical fallacies

1. INTRODUCTION

The purpose of *human reliability assessment* (HRA) is to determine how likely it is that operators will incorrectly perform one or more of the required safety-critical tasks in response to some event. This stems from the *event tree* and *failure tree* (ET & FT) structure of the *probabilistic safety assessment and management* (PSAM), which is the general way to represent the event scenario probabilistically. Therefore, HRA can be seen as a way to expand ET & FT nodes representing internal (technological) or external (environmental) events with *human failure events* (HFE), corresponding to *human error probabilities* (HEP) [1].
Typically, each node in the trees is characterized by the specific conditions under which a given facility may succeed or fail. Since the hardware and software safety systems are usually independent by design, simple tree graph structures fit the current configuration of at-risk facilities such as nuclear power plants (NPP), chemical and aerospace systems with sufficient confidence [2].

However, this does not apply to human activity and HRAs, since the human actions (HA) are not only independent of each other, but even depend simultaneously and holistically on the internal and external events, organization of work and the person himself. This means that they vary greatly depending on the specific situation, determined by the sequence and importance of previous and future symptoms (events, goals, parameters, functions, resources, transitions and actions), i.e. on the **context of entire system**, consisting of the **object**, as a *nuclear power plant - NPP*, and the **subject** recognizing it, as a crew of operators, in a given situation. Context is qualitatively & quantitatively determined by the *Violated Objectively (VO*) or *Objectively (O)* non-violated symptoms of the object from their *Subjectively (S)* recognized shifted images by the subject in a given situation. "*Context, as the description of those aspects of the system state that makes the operator conclude that his behavior is correct, is one of the most important issues for understanding and tackling errors of commission"* [3].

Considering errors such as **EOO** (*error of omission*) and **EOC** (*error of commission*) as complementary categories for the overall characterization of unsafe acts leads to certain problems with their distinction into an HFE taxonomy, as shown in Figure 1. These are related to the fact that EOO and EOC can occur both in the cognitive/diagnostic part of human action and in its executive part. Thus, as Hollnagel [1] argues, EOC can be viewed as both a cause (genotype) and a manifestation/effect (phenotype). However, this problem does not arise, if we consider only *logical HFE*, such as the conjunction and disjunction fallacies, which are primarily associated with cognitive and decision-making processes.

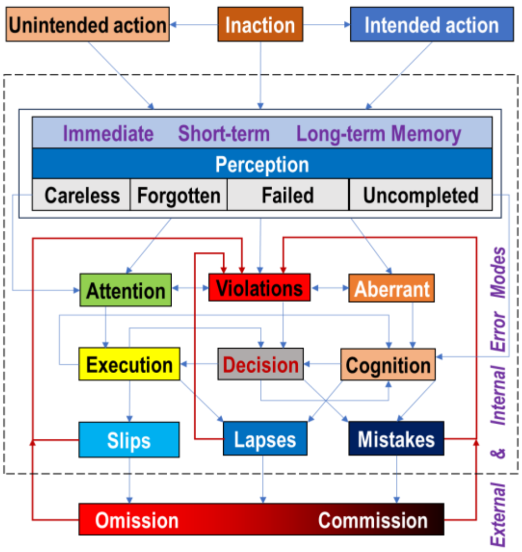

Figure 1. HRA taxonomy of a logical EOO/EOC HFE, in the glocal system context [4].

Therefore, they can be defined as follows:

- An **EOO** is a partial or complete disregard, omission, or neglect of something you should have done. A **logical EOO** is the failure to recognize an opportunity or decide to take some action necessary to achieve a desired goal.
- A **logical EOC** is recognizing or deciding to perform an unrelated action that impedes or creates obstacles to the achievement of the goal, i.e. *"interventions of operators that are not required from the system point of view and aggravate the scenario evolution."* The EOC are only from a mental point of view, *"not errors in the sense that a person failed to perform correctly"* [3]. Operators making an EOC usually perform correctly from their current understanding of the system context defined as an ensemble of past, current and future system states.

Any useful HRA method, not only for the PSAM but also for understanding thought processes, must successfully combine psychological description (*"descriptive correctness"*) with mathematical modeling and systematic classification with deterministic-stochastic evaluation (*"methodological correctness"*). Theoretical knowledge is often insufficient and therefore must be supplemented with imagination and the use of heuristic methods for simulation and data collection for correct statistical processing (*"statistic correctness"*) [3].

Data usually combines two types of sources: *empirical evidence* and *expert judgment* sources [2]. The use of each of these sources requires knowledge of at least two specific contexts: 1) the context of the action being evaluated, and 2) the context of the evaluation itself, to account for the relevant logical fallacies and ambiguities inherent in human cognition and judgment. Therefore, using a symptom-based context-sensitive HRA method, rather than an HRA method mixing cause and effect together (via Performance Shaping Factors - PSF), halves the search for contextual influences and dependencies between these at least two contexts.

Therefore, this paper attempts to explain and assess the probabilities of misjudgment in an ambiguous and comparative context (e.g. Linda problem, Prisoners' dilemma, Ellsberg's and Machina's paradoxes) with a view to properly evaluating and classifying HFEs by the symptom-based context-sensitive PET method. Particular attention is paid to the logical fallacies of conjunction and disjunction made by people in their work with at least two groups of symptoms that are perceived by them as *significant* and *insignificant* for the context (object and subject in situation) in which they have to make the *logical judgment* of conjunction or disjunction.

Accurately assessing the statistical probability of a misjudgment requires calculating the ratio of the number of logical fallacies to the number of people interviewed in surveys, for example, the conjunction error in the Linda problem [5].

The probability of misjudgment is calculated by counting the alternatives/trajectories of the system contexts. Statistically concordant estimates of these fallacies between surveys and those obtained with the PET method (observed information vs. predictive models) indicate that such context quantification is reasonable and useful. Therefore, they can be used both for verification and validation of the PET method and for collecting data on the probability of misjudgment in simulator training of NPP operators [6].

2. WHOLENESS AND IMPLICATE ORDER OF COGNITION AND JUDGMENT

According to the ideas of David Bohm about implicate and enfolded order, the processes on the macro- and micro-levels are analogous and isomorphic and obey the same rules [7]. He also notes that consciousness can be described in terms of a series of moments, i.e. *"one moment gives rise to the next, in which context that was previously implicate is now explicate while the previous explicate content has become implicate"* [7]. Consciousness is an interchange process with feedback that results in a growing accumulation of understanding.

The model of the context-cognition recurrent interaction, on the surface/macro (or in-depth/micro) level, should represent a network with cognition sub-processes as nodes, where edges/links enter a sub-process domain from different directions (other sub-processes) to set up interference patterns. As factorial experiments and empirical task investigations suggest, there exist sub-processes in the cognitive, decision-making process. They explain that context images should be an interference of context factors or symptoms (cues, signals, signs or symbols) and should be perceived as a hologram. The context-cognition interaction is performed by information processing and exchange of "units" of information (bits). A macroscopic information entropy parameter (context) has to connect and integrate sub-processes of cognition. Statistical mechanics demonstrates that energy entropy is governed by probability, and for the information entropy a macroscopic parameter is *context probability* (CP) or *'contexture'*. The term 'contexture'[1] is coined by analogy with 'temperature' as a potential to exchange information vs. heat, to show the similarity between information entropy and energy entropy as a distribution of information vs. energy.

Failure to complete the cognition and judgment process in ambiguous and comparative contexts is due to limited, delayed and stochastic abilities to encode information in *working memory*. This, in turn, limits subsequent access to this information in *long-term memory*, where the *"individual judgment processors"* are stored and used through a unified probabilistic thinking logic.

3. SIMULATOR DATA USE FOR PSA PURPOSES

The data collection requirements for HRA purposes are much the same as for training [8].

HRA seeks information about the HEP and their causes. There is usually not enough data on HFE in a given scenario to draw a solid conclusion about the HEP. In a multi-unit NPP, the number of crews participating in the training for a given scenario is 6 per unit. If the NPP has 2 units (like Kozloduy NPP) then the total number of participating teams will be 12 and the lowest HEP value that can be obtained in one training session for data collection will be:

$$n_e / (n_s * n_c) = 1 / (1*12) = 0{,}083 \quad (1)$$

where: $\boldsymbol{n_e}$ - *number of erroneous actions*, $\boldsymbol{n_s}$ - *number of scenarios played*, $\boldsymbol{n_c}$ - *number of crews*.

HEP values can vary from 1,0 to 1,0E-03, and for EOC they are even smaller! Therefore, in order to be able to collect enough data for EOC, for example with HEP 1,0E-03, we have the following three options [9]:

1) *Extensive training statistics* - more training sessions on a given scenario (*economically unreasonable*);
2) *Shared statistics* based on joint agreement and data mining framework (*practically unfeasible*);

[1] The term context could be kept for the regular use.

3) *Intensive experimental statistics* to give interpretation and evaluation of extraordinary circumstances. As highlighted by Spurgin & Petkov (2005) this is *theoretically unproven*. Therefore, without solid theoretical and experimental evidence, it will be difficult to obtain sufficient data on EOC.

Our goal here is to try to justify how we can theoretically increase the capabilities of experimental statistics based on nuclear facility simulators to obtain sufficient data to extract EOCs that may or may not lead to core damage. In addition, minor errors may occur even more frequently, such as missing a step, delays, or confusion in the order of steps (EOO) in symptom-based emergency procedures, which should be distinguished and subtracted from the EOC statistics. All these errors are often corrected and recovered later. But there is no guarantee that minor errors can always be recovered in a different context. The impact of these errors is significant not only from a training perspective, as crews deviate from the proper operation of the unit and the entire facility. In addition, these errors can be interpreted as violations that not only affect the development of the accident, but also change the processes of cognition and decision-making and are contextual possibilities for additional alternatives for the occurrence of EOC (as shown by red dark arcs in Figure 1).

Data through the full-scale NPP simulator is collected during the main control room crew's responses to various normal, emergency or abnormal scenarios. Operators follow symptom-oriented procedures and some of the steps can be repeated in a few scenarios, of which there are about 10 in one training session. The interpretation of errors follows a certain taxonomy, but the distribution of errors is not the same for every scenario because it follows the assumed context dependence. The PET method for qualitative and quantitative context evaluation allows the scenario to be adapted to worsen, equalize or improve its context played out on the simulator with the same action, i.e. $n_s > 1$. In this way, it is possible to model incomplete or violated cognition, which increases the probability of EOC.

Figure 2 is an illustration of the context influence on EOC, EOO, and recovery [9].

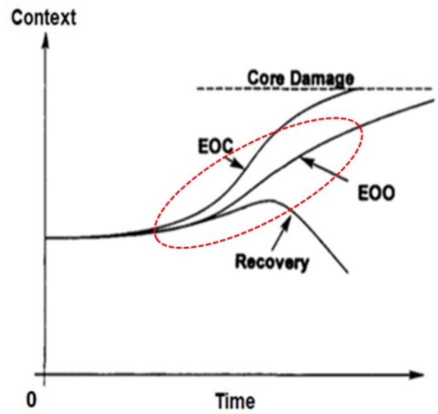

Figure 2. Illustration of the context influence on EOC, EOO, and recovery [9].

4. WRONG LOGIC OF JUDGMENT IN COMPARATIVE AND AMBIGUOUS CONTEXT

4.1. Extended Description of the Linda Problem for Conjunction and Disjunction Fallacy

This application estimates conjunctive and disjunctive misjudgments with a unified framework of the PET method. A modified version of the most famous demonstration of the conjunction fallacy, a *Linda problem*, adapted for descriptive purposes and incorporating the disjunction fallacy [10].

"Linda is 31 years old (31), single (S), outspoken (O), and very bright (B). She majored in philosophy (P). As a student, she was deeply concerned with issues of discrimination (D) and *social justice (J), and also participated in antinuclear demonstrations (A)." Participants were asked to rank various statements about Linda "by their probability". Two of these statements were:*

- *Linda is a bank teller (T);*
- *Linda is a feminist (F);*
- *Linda is a bank teller and active in the feminist movement (T and F);*
- *Linda is a bank teller or active in the feminist movement, or both (T or F).*

4.2. Definitions of Conjunction and Disjunction Fallacy by the PET Context Quantification Procedure

The PET method makes it possible to identify and evaluate the probabilities of conjunction and disjunction fallacies when modeling them in a comparative and ambiguous context and assessing varying *contexture* (CP) and *cognitive error probability* (CEP). These context-sensitive assessments allow uncertainty to be rationally interpreted within the error probability margins. In this application of the PET method, the extended Linda problem of a misjudgment is solved by comparing the contexts of individual constituents and algebraic expressions obtained using the conjunction fallacy (**fallacy C**) logic [6]:

- *single conjunction fallacy*: $CP(T) \leq CP(T \land F)$ **OR** $CP(F) \leq CP(T \land F)$
- *double conjunction fallacy*: $CP(T) \leq CP(T \land F)$ **AND** $CP(F) \leq CP(T \land F)$.

The disjunction fallacy (**fallacy D**) could be described in a similar way:

- *single disjunction fallacy*: $CP(T) \geq CP(T \lor F)$ **OR** $CP(F) \geq CP(T \lor F)$
- *double disjunction fallacy*: $CP(T) \geq CP(T \lor F)$ **AND** $CP(F) \geq CP(T \lor F)$.

4.3. Contexture of Fallacy Using Stepwise Models of Cognitive Entanglement

The PET method uses stepwise models of cognitive processes to recognize symptoms during a person's judgments and actions [11]. These stages are similar to *Donders' stage durations* [12], but can be used in different orders, combining sequential and parallel recognition of symptoms. To obtain estimates of the probabilities of conjunction and disjunction fallacies, it is necessary to use at least two simple types of symptom recognition, consisting of different parallel and sequential stages of violated and/or incomplete cognitive processes.

They are shown in Table 1 respectively. It shows the Excel table for this context without violations (***abc_de_fg_hi***) = (**000_00_00_00**) for which the values exactly corresponding to the rules of:

- **conjunction** $CP(T \land F) = CP(F)*CP(T)=0,5*0,5=0,25$ (after 2^{nd} Step) and
- **disjunction** $CP(T \lor F) = CP(F) + CP(T)=0,5 + 0,5=1$ (after 5^{th} Step).

Table 1: Calculating the Contexture (CP) of the recognition processes of (T), (F), (TΛF) and (TVF) models without (0) or with (≥1) violations for all three groups of symptoms, only the "0" models are shown here.

Cognition	Constituents truth	Context Factors and Conditions (Goups of Symptoms)									CP
	000_00_00_00	V_(T Λ F) or V_(T V F)			V_T			V_F			
	Steps	VO	O	S	VO	O	S	VO	O	S	
T	1	a=0	0	0	d=0	1	0	e=0	0	0	1
	2	a=0	0	0	d=0	1	1	e=0	0	0	0,5
F	1	b=0	0	0	f=0	0	0	g=0	1	0	1
	2	b=0	0	0	f=0	0	0	g=0	1	1	0,5
Cognition	Conjunction	Context Factors and Conditions (Goups of Symptoms)									CP
	000_00_00_00	Conjunction V_(T Λ F)			V_T			V_F			
	Steps	VO	O	S	VO	O	S	VO	O	S	
(T Λ F)	1	c=0	0	0	h=0	1	0	i=0	1	0	1
	2	c=0	0	0	h=0	1	1	i=0	1	1	0,25
Cognition	Disjunction	Context Factors and Conditions (Goups of Symptoms)									CP
	000_00_00_00	Disjunction V_(T V F)			V_T			V_F			
	Steps	VO	O	S	VO	O	S	VO	O	S	
(T V F)	1	c=0	0	0	h=0	1	0	i=0	0	0	1
	2	c=0	0	0	h=0	1	1	i=0	0	0	0,5
	3	c=0	0	0	h=0	1	1	i=0	1	0	1,5
	4	c=0	0	0	h=0	1	1	i=0	1	1	0,5
	5	c=0	0	0	h=0	1	1	i=0	1	1	1

It is assumed that the process of recognizing the two constituents (T & F groups of symptoms), in order not to make a conjunction fallacy (single or double), can be completed practically in two parallel steps of their recognition. However, the final completion of cognition will occur in two steps when all symptoms are recognized [13]. To avoid the fallacy of disjunction (single or double), the process of recognizing these two groups of symptoms (T and F) can be carried out in five successive steps. However, for their complete utilization, additional steps are required to completely ignore them [13], [14]. The time it takes to fully recognize a symptom is comparable to the time it takes to disregard it. It is the process of erasing information from a subject's *short-term working memory* [15].

5. LOGIC OF THINKING ABOUT THE EXTENDED LINDA PROBLEM: MODELS AND RESULTS OF THE PERFORMANCE EVALUATION OF TEAMWORK METHOD

5.1. Relevant and Irrelevant Information for Misjudgment in Ambiguous and Comparative Context

An odds ratio satisfying the conjunction rule will be obtained if the respondent is not given any prior information about Linda's problem (31, S, O, B, P, D, J, A). There is then no ambiguity and the respondent does not use these facts/symptoms and the question comes down to a simple comparison between (T) and (F) and the application of the classical rule of conjunction.

When the question is asked in an ambiguous and comparative context, the comparison of (T∧F)/(TVF) with (T) and (F) is also based on a priori information about Linda. These are facts of the past that carry relevant information for comparison, but because they may have already changed, they give rise to uncertainty (**Violation with Relevant information, VR**). Therefore, they can be considered as a violation in the group of symptoms (T) and (F) that disorder the cognition/judgment processes of the alternatives (T), (F) and (T∧F)/(TVF). This is a violation that may only be related to one of symptoms (T) or (F), but may affect one, two or all three judgments (T), (F) and (T∧F). It means that ***d***, ***e***, ***f***, ***g***, ***h*** and ***i*** could be **0** or **1**, but ***a***=***b***=***c***=***0***. The coding of all context trajectories starts with a single non-violated variant (**000_00_00_00**) and may vary for a total of **3^3=27** context alternatives. Indifferent (noisy information) facts for Linda only introduce confusion and uncertainty into the application of the conjunction rule (**Violation with Indifferent information, VI**). Therefore, this violation can appear only in the alternative ((T∧F)/(TVF), i.e., ***c*** = ***1***), but not in a trivial comparison of alternatives (T, i.e., ***a*** = ***0***) and (F, i.e., ***b*** = ***0***), since it does not provide relevant information about them. The difference between columns of ***symmetrical cases***, [P(T)=P(F)=0,5], and ***asymmetrical cases***, [P(T)≠P(F)], is that the number of objectively occurring and recognized symptoms for (T), (F) and (T∧F)/(TVF) in V_F and V_T is different, but the distribution of violations (***a***, ***b***, ***c***, ***d***, ***e***, ***f***, ***g***, ***h***, ***i***) is the same.

5.2. Conjunction in Ambiguous and Comparative Context

As can be seen in Tables 2 (2-1 & 2-2), the conjunction fallacy model is only one and it is easy to make comparisons with the results of the article by Tversky and Kahneman and their followers [5].

Table 2-1. The Only One Fallacy C Model of Conjunction Fallacy.

True to False Cognition	Conjunction Cog	Context Factors and Conditions (groups of symptoms)									CP
	000_00_00_00	V_(T ∧ F)			V_(F)			V_(T)			
	Steps	VO	O	S	VO	O	S	VO	O	S	
(T)	1	0	0	0	0	1	0	0	0	0	1
	2	0	0	0	0	1	1	0	0	0	0,5
(F)	1	0	0	0	0	0	0	0	1	0	1
	2	0	0	0	0	0	0	0	1	1	0,5
(T ∧ F)	1	0	0	0	0	1	0	0	1	0	1
	2	0	0	0	0	1	1	0	1	1	0,25

Tables 2-2 shows the results of 12 cases for contextual trajectories of a conjunction fallacy model.

Table 2-2. The Conjunction Fallacy C

Variant	sym_000_	sym_100_	sym_010_	sym_001_	sym_110_	sym_101_	sym_011_	sym_111_	asym_2_1	asym_1_2	asym_3_1	asym_1_3
double	2	0	0	19	0	6	6	2	14	14	11	11
single	8	6	6	4	0	15	15	8	11	11	14	14
success	17	21	21	4	27	6	6	17	2	2	2	2
Fallacy C	37,04%	22,22%	22,22%	85,19%	0,00%	77,78%	77,78%	37,04%	92,59%	92,59%	92,59%	92,59%

The results show 12 cases of *misjudgment in comparative and ambiguous contexts*, each with 27 contextual trajectories [6]. The cases of *conjunction fallacies* (fallacy C) shown in **green** (**0,37037**, **dark** or **light**) are very close to the *"correct answer to the short version"* result obtained by Tversky & Kahneman [5] for the Linda problem (**36%** failure). The case for *conjunction fallacies* (fallacy C) colored **yellow 0,851852** and **pale red 0,925296** practically match (coincides or slightly exceed the statistical result) the "stark version" (about **"85% to 90%"**) obtained by Tversky and Kahneman and their followers on the Linda problem [5].

5.3. Invariance of Disjunction in Ambiguous and Comparative Context

Unfortunately, this is not the case when compared to the disjunction fallacy results.

As can be seen from Tables 3 (3-1 & 3-2), (4-1 & 4-2) and Table 5 (5-1 & 5-2), the models could be invariant: *Disjunction Fallacy of Full Cognition* (D_FC), *Disjunction Fallacy of Relay Cognition* (D_RC), *Disjunction Fallacy of Serial Cognition* (D_SC) and different results can be obtained.

Table 3-1. Invariant of Disjunction Fallacy: D_FC Model.

True to False Cognition	Disjunction Full Cog	Context Factors and Conditions (groups of symptoms)									CP
	000_00_00_00	V_(T V F)			V_F			V_T			
	Steps	VO	O	S	VO	O	S	VO	O	S	
(T)	1	0	0	0	0	**1**	0	0	0	0	1
	2	0	0	0	0	**1**	**1**	0	0	0	**0,5**
(F)	1	0	0	0	0	0	0	0	**1**	0	1
	2	0	0	0	0	0	0	0	**1**	**1**	**0,5**
(T V F)	1	0	0	0	0	1	0	0	0	0	1
	2	0	0	0	0	1	1	0	0	0	0,5
	3	0	0	0	0	1	1	0	1	0	1,5
	4	0	0	0	0	1	1	0	1	1	**0,5**
	5	0	0	0	0	1	1	0	1	1	**1**

Table 3-2. D_FC Disjunction Fallacy.

Variant	sym_000_	sym_100_	sym_010_	sym_001_	sym_110_	sym_101_	sym_011_	sym_111_	asym_2_1	asym_1_2	asym_3_1	asym_1_3
double	17	21	21	1	27	3	3	17	10	17	5	18
single	7	6	6	4	0	18	18	8	13	8	19	9
success	3	0	0	22	0	6	6	2	4	2	3	0
Fallacy D_FC	88,89%	100,00%	100,00%	18,52%	100,00%	77,78%	77,78%	92,59%	85,19%	92,59%	88,89%	100,00%

Table 4-1. Invariant of Disjunction Fallacy: D_RC Model.

True to False Cognition	Disjunction Relay Cog	Context Factors and Conditions (groups of symptoms)									CP
	000_00_00_00	V_(T V F)			V_F			V_T			
	Steps	VO	O	S	VO	O	S	VO	O	S	
(T)	1	0	0	0	0	**1**	0	0	0	0	1
	2	0	0	0	0	**1**	**1**	0	0	0	**0,5**
(F)	1	0	0	0	0	0	0	0	**1**	0	1
	2	0	0	0	0	0	0	0	**1**	**1**	**0,5**
(T V F)	1	0	0	0	0	1	0	0	0	0	1
	2	0	0	0	0	1	1	0	0	0	0,5
	3	0	0	0	0	1	1	0	1	0	1,5
	4	0	0	0	0	0	0	0	1	1	**0,5**
	5	0	0	0	0	0	0	0	1	1	**1**

Table 4-2. D_RC Disjunction Fallacy.

Variant	sym_000_	sym_100_	sym_010_	sym_001_	sym_110_	sym_101_	sym_011_	sym_111_	asym_2_1	asym_1_2	asym_3_1	asym_1_3
double	13	15	15	1	22	3	3	13	8	16	3	15
single	4	9	9	4	4	12	12	4	8	5	12	9
success	10	3	3	22	1	12	12	10	11	6	12	3
Fallacy D_RC	62,96%	88,89%	88,89%	18,52%	96,30%	55,56%	55,56%	62,96%	59,26%	77,78%	55,56%	88,89%

Table 5-1. Invariant of Disjunction Fallacy: D_SC Model

True to False Cognition	Disjunction Serial Cog	Context Factors and Conditions (groups of symptoms)									CP
	000_00_00_00	V_(T V F)			V_F			V_T			
	Steps	VO	O	S	VO	O	S	VO	O	S	
(T)	1	0	0	0	0	1	0	0	0	0	1
	2	0	0	0	0	1	1	0	0	0	0,5
(F)	1	0	0	0	0	0	0	0	1	0	1
	2	0	0	0	0	0	0	0	1	1	0,5
(T V F)	1	0	0	0	0	1	0	0	0	0	1
	2	0	0	0	0	1	1	0	0	0	0,5
	3	0	0	0	0	0	0	0	1	0	1,5
	4	0	0	0	0	0	0	0	1	1	0,5
	5	0	0	0	0	0	0	0	0	0	1

Table 5-2. D_SC Disjunction Fallacy.

Variant	sym_000_	sym_100_	sym_010_	sym_001_	sym_110_	sym_101_	sym_011_	sym_111_	asym_2_1	asym_1_2	asym_3_1	asym_1_3
double	14	18	18	1	27	3	3	14	8	17	3	17
single	7	9	9	4	0	15	15	8	11	8	15	8
success	6	0	0	22	0	9	9	5	8	2	9	2
Fallacy D_SC	77,78%	100,00%	100,00%	18,52%	100,00%	66,67%	66,67%	81,48%	70,37%	92,59%	66,67%	92,59%

5.4. Results of Symmetrical and Asymmetrical Models for Conjunction and Disjunction Fallacy

Table 6 shows the codes, number of cases and the probability of conjunction and disjunction fallacy (in %) for contextual trajectories of 8 symmetrical cases, P(T)=P(F)=0,5. Figure 3 shows statistical predictions for logical fallacy probabilities of symmetrical cases.

Table 6: The results of 8 symmetrical contextual trajectories models of conjunction and disjunction fallacy.

Symmetrical Case	000_sym 1	100_sym 2	010_sym 3	**001_sym 4**	110_sym 5	101_sym 6	011_sym 7	111_sym 8
Conjunction	37,0%	22,2%	22,2%	85,2%	0,0%	77,8%	77,8%	37,0%
Disjunction_FC	92,6%	100,0%	100,0%	18,5%	100,0%	77,8%	77,8%	92,6%
Disjunction_RC	63,0%	88,9%	88,9%	18,5%	96,3%	55,6%	55,6%	63,0%
Disjunction_SC	77,8%	100,0%	100,0%	18,5%	100,0%	66,7%	66,7%	81,5%

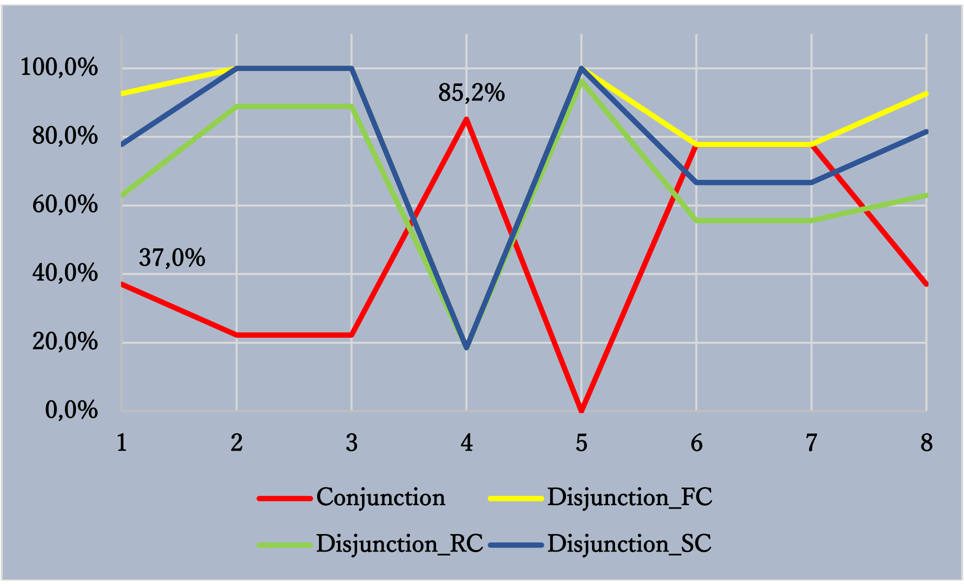

Figure 3. Statistical Predictions for Logical Fallacy Probabilities of Symmetrical Cases

Table 7 shows the codes, number of variants and the probability of conjunction and disjunction fallacy (in %) for contextual trajectories of 4 asymmetrical cases, where P(T)≠P(F). Figure 4 shows statistical predictions for logical fallacy probabilities of asymmetrical cases.

Table 7: The results of one symmetrical and 4 asymmetrical contextual trajectories models of conjunction and disjunction fallacy.

Asymmetrical Cases	**001_sym 4**	001_asym_2_1 I	001_asym_1_2 II	001_asym_3_1 III	001_asym_1_3 IV
Conjunction	85,2%	92,6%	92,6%	92,6%	92,6%
Disjunction_FC	18,5%	85,2%	92,6%	81,5%	100,0%
Disjunction_RC	18,5%	59,3%	77,8%	55,6%	88,9%
Disjunction_SC	18,5%	70,4%	92,6%	66,7%	92,6%

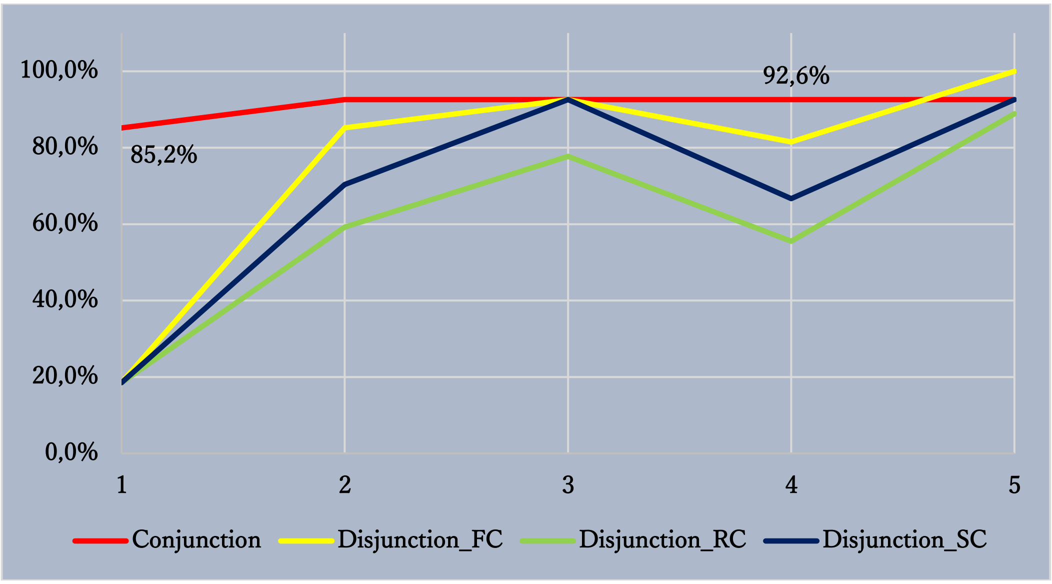

Figure 4. Statistical Predictions for Logical Fallacy Probabilities of Asymmetrical Cases

6. CONCLUSION

1. Failure to complete the cognition and judgment process in ambiguous and comparative contexts is due to limited, delayed and stochastic abilities to encode information in *working memory*. This, in turn, limits subsequent access to this information in *long-term memory*, where the "individual judgment processors" are stored and used through a unified probabilistic thinking logic.

2. Avoiding cognitive biased misjudgment is primarily concerned with limiting the effects of delayed or violated symptom recognition through careful qualification, grouping and quantification of symptoms. The modeling, explanation, and solution of the *conjunction-disjunction fallacy* (extended Linda problem) allows us to present a heuristic and rational methodology for interpreting many experimental psychological studies of problems, dilemmas and paradoxes.

3. The explicative-implicative estimates of conjunction and disjunction fallacies can be thought of as estimates of the predicted statistical probabilities of such inherent human logical biases. Qualitative and especially quantitative assessment of context, for example using an HRA method such as the PET technique, can be used as a tool to predict the probability of logical fallacies.

4. Following the severe nuclear accidents, full-scale nuclear facility simulators are increasingly used to simulate rare and unexpected scenarios with multiple failures and violations. EOCs of logical conjunctions and disjunctions, although unlikely, may arise during routine operator training on simulators.

5. Therefore, the use of models and methods such as PET to determine and assess the context of the accident through simulator data can increase the confidence in HRA techniques and generate useful EOC-specific databases.

6. The usefulness of simulators for training purposes is undeniable. However, their usefulness for HRA purposes has been challenged by favoring the use of an HRA method that relies on a common database

rather than collecting and evaluating site-specific HEPs based on crew responses to simulated scenarios.

7. Simulators have limitations, but they are the best invented experimental tool for integration, coherence and coordination of processes in the system: human, technology, organization, and even environment.

8. NPP simulators equipped with modems recording audio and video data collection systems can be considered as the most important laboratories for effective research into the safety and efficiency of the human-technology-organization system.

9. The HRA methods have even more shortcomings and without the help of representative simulators' data, it would be difficult for the HRA methods to be corrected, improved, verified and validated [16].

10. In addition to addressing the problems of logical misjudgment, this paper aims to explore and improve the PET method, since without data feedback, no HRA method can achieve the required quality and applicability. Making more rational risk-informed decisions despite human biases may be one of the most important applications of the symptom-based contextualization for probabilistic safety assessment and management.

References

[1] Hollnagel E. Looking for errors of omission and commission or The Hunting of the Snark revisited. Reliability Engineering and System Safety, 68, 135-145, 2000.
[2] Greco S.F., Podofillini L. and Dang V. N. A Bayesian two-stage approach to integrate simulator data and expert judgment in human error probability estimation. Safety Science, 159, 2023, 106009.
[3] Strater O., Dang V., Kaufer, B. and Daniels, A. On the way to assess errors of commission. Reliability Engineering and System Safety, 83, 129–138, 2004.
[4] Petkov, G., Holistic Context Analysis for Dynamic HRA and Accident Management, IAEAJ4-TM-55244, IAEA TM on the Development of the Safety Report on Human Reliability Assessment for Nuclear Installations, Vienna, Austria, 13-17 November, 2017 P7_Petkov_IAEA_J4-TM-55244.
[5] Khaneman, D. Thinking, Fast and Slow, Farrar, Straus and Girou, 2011.
[6] Petkov G. and Petkov I. Context of Natural and Artificial Intelligence Misjudgment in Nuclear Technology, Proceedings of the PSAM Topical Conference 2023, 23-25 October, University of Illinois Urbana Champain, (virtual).
[7] Bohm, D. Wholeness and the Implicate Order, Routledge, 2002.
[8] Spurgin, A. J., Bareith, A. and Karsa. Z. Simulator Data Collection Requirements for HRA, Proceedings of the PSAM 7 Conference, Berlin, Germany, 14-18 June, 2004.
[9] Spurgin A. J. and Petkov, G. I. Power Plant Simulators: Data Collection and Insights from Data, Journal of Safety and Reliability 'Human Performance Data Collection Edition', Vol. 26, No. 2, 47-67, Summer 2006, ISSN 0961-7353.
[10] Nilsson, H., Rieskamp, J., & Jenny, M.A. Exploring the overestimating of conjunctive probabilities. Frontiers in Psychology, 4, 1-12, 2013.
[11] Petkov, G. Symptom-based context quantification for dynamic accident analysis, Safety Science Journal, 2018, https://doi.org/10.1016/j.ssci.2018.02.027
[12] Donders, F. Die Schnelligkeit Psychischer Processe, Archiv fur Anatomie and Physiologie, 657-681, 1868. (In W.G. Koster Ed. Attention and Performance II (pp.176-315), Elsevier/North-Holland.
[13] Petkov, G. Enhancing the plant's capability for design extension conditions based on best estimate context evaluation of human performance in ATWS events, ICONE24 2016 Conference, ICONE24-60707, Charlotte, NC, USA, June 26-30, 2016.
[14] Rodin, M.J. Who is memorable to whom? A study of cognitive regard, Social Cognition, 5, 144-165, 1987.
[15] Cowan, N. *The magical number 4 in short-term memory: A reconsideration of mental storage capacity.* Behavioral and Brain Sciences 24: 87–185, 2001.
[16] Spurgin, A. and Petkov, G. Advanced Simulator Data Mining for Operators' Performance Assessment, in book Intelligent Data Mining: Techniques and Applications, Physica Verlag, pp.543-572, 2005.