

文章编号: 1001-2486(2008)01-0073-05

基于认知模型的人为差错分类方法*

孙志强, 史秀建, 李欣欣, 谢红卫

(国防科技大学 机电工程与自动化学院, 湖南长沙 410073)

摘要: 为了保证人为差错分类的全面性和一致性, 在详细分析现有人为差错分类方法的基础上, 提出了一种新的人为差错分类框架。该框架以“执行差错/忽略”为起点进行细化, 得到五种基本的人为差错类别。以此作为人为差错原语, 结合认知模型, 分析了不同认知阶段可能出现的人为差错模式。实例证明, 在这种分类方法指导下来分析人为差错, 对提高人为差错分类结果的全面性和一致性有着重要意义。

关键词: 人为差错; 差错类别; 执行差错; 忽略差错; 认知模型; 系统安全

中图分类号: TB114.3 **文献标识码:** A

A Framework for Classifying Human Errors Based on Cognitive Model

SUN Zhi-qiang, SHI Xiu-jian, LI Xin-xin, XIE Hong-wei

(College of Mechatronics Engineering and Automation, National Univ. of Defense Technology, Changsha 410073, China)

Abstract: A new framework for classifying human errors is presented on basis of the existing frameworks. The human error taxonomies derived from this framework are more comprehensive and consistent. In this framework, the seminal framework of “Commission/Omission” is branched out, and five basic human error types are derived. The five human error types can be used as primitives and combined with cognitive model in order to identify the human error modes at various cognitive modules. The process of human error identification is illustrated by an artificial example. The results show that this framework is effective.

Key words: human error; error taxonomy; commission error; omission error; cognitive model; system safety

“智者千虑, 必有一失”, 在某种意义上, 有人干预的地方都有可能出现人为差错。目前, 绝大部分武器装备在使用过程中都离不开人为干预和支持。人为差错能够直接影响到装备的使用安全和效能发挥。因此, 必须充分重视人因可靠性分析(HRA, Human Reliability Analysis)工作, 分析可能的人为差错模式并设计相关规避措施。HRA过程可以分为四个阶段: 人为差错辨识与描述、人为差错发生概率计算、差错后果分析和规避措施设计。从某种意义上讲, 人为差错辨识是HRA的基础, 全面辨识出所有可能发生的人为差错模式, 能够在很大程度上提高HRA结果的可信性, 拓展HRA结果的意义^[1]。

与普通的物理部件不同, 人的差错模式表现出多样性。如果仅仅依赖分析人员的领域知识和主观判断来辨识人为差错, 那么, 一方面很难保证辨识结果的全面性, 另一方面对辨识结果的描述可能也各有不同, 甚至出现本质上相同的差错却被冠以不同“术语”的情况, 降低了辨识结果的可读性。因此, 从保证人为差错辨识结果的全面性和一致性的角度出发, 设计一种能够涵盖所有差错模式的人为差错分类基本框架, 用于指导人为差错辨识过程, 是一件非常有必要也是非常有意义的工作。

1 人为差错分类方法概述

目前, 常见的人为差错分类方法有以下几种: 执行差错/忽略(Commission/Omission)、失误/差错/错误(Slip/Lapse/Mistake)和基于认知模型的分阶段人为差错分类方法。其中, 执行差错/忽略分类框架着重从差错的表现形式着手, 分析差错动作的外在形式; 而失误/差错/错误这种分类框架则是从“意图”出发, 分析差错背后隐含的操作意图。分阶段的人为差错分类方法强调将认知过程分为几个阶段, 然后分

* 收稿日期: 2007-09-30

基金项目: 国家部委资助项目(51319040201)

作者简介: 孙志强(1978—), 男, 博士生。

析不同认知阶段可能发生的差错。

1.1 执行差错/忽略

“执行差错/忽略”是 THERP(A Technique for Human Error Rate Prediction) 方法所使用的人为差错分类方法^[2]。THERP 方法是一种人因可靠性分析方法, 1983 年最终成型, 在核工业中有着较为广泛的应用。这种分类方法描述了两种不同形式的差错, 其中, 执行差错表示“针对某个任务, 采取了某个执行动作, 但这个动作是错误的”, 这包括两方面的含义, 其一, 当任务要求是不执行任何动作时, 那么此时若采取了动作就属于执行差错; 其二, 任务要求采取相关动作时, 此时若采取了与要求不相符合的动作也属于执行差错。忽略则表示“在应该执行相关动作的情况下, 没有采取任何动作”。

可以看出, 这种分类方法比较粗糙。同时, 它又是一种非常全面的分类方法, 能够涵盖所有的人为差错模式。考虑到这一点, 可以以此为起点, 通过细化分析, 得到新的人为差错分类方法。

1.2 失误/差错/错误

这种分类方法是 Reason 于 1990 年提出的^[3], 与“执行差错/忽略”不同, 它着重从“意图”的角度出发, 分析不同的意图可能导致的人为差错。所谓意图, 指的是操作人员在执行某个动作之前, 所进行的决策过程及决策结果。

在这种分类方法中, 失误表示“意图正确, 但采取了错误的动作”, 差错表示“意图正确, 但在应该采取动作的时候没有采取任何动作”, 错误则表示“意图是错误的, 执行的动作也是错误的”。比如, 当核电厂出现紧急情况时, 此时应该关闭阀门 A, 操作人员如果做出了正确的判断, 但关闭的却是阀门 B, 那么此时的差错就是失误; 如果此时操作人员没有采取任何动作, 则属于差错; 而操作人员如果做出的判断是关闭阀门 B, 同时也关闭了阀门 B, 那么此时的差错就属于错误。

表面上看起来, 由于考虑了差错出现的意图因素, 这种分类方法似乎比前一种分类方法更为深入, 但是, 实际上, 在人为差错分析过程中, 很多情况下可能很难确切知道操作人员的意图究竟是什么, 能够看到的仅仅是差错的表现形式。在前面的示例中, 如果操作人员拉下了阀门 B, 对他的决策过程是否正确可能并不知情。因此, 这种分类方法相对来说更加适合于事故分析, 当事故发生后, 通过调查相关人员, 回溯分析当时的分析决策过程。已经有很多研究人员利用这种分类框架对飞行差错、医疗差错进行分类^[6-8]。

1.3 基于认知模型的分阶段的人为差错分类方法

以上这两种分类方法都是从任务执行的最后阶段出发来分析人为差错的, 基本上不考虑差错发生机理。对于简单的日常操作任务来说, 以上分类方法可能足以胜任了。随着工业生产的进步, 人在生产过程中的角色不再是纯粹的“操作员”, 正逐渐向“监控员”转化。在这种情况下, 任务执行过程更多的是需要操作人员根据获取的信息及时进行判断与决策。这样一来, 能够导致出现人为差错的因素不仅仅在于“执行”这一阶段, “信息获取”和“决策”等阶段的差错同样可以导致严重后果。因此, 一部分 HRA 方法倾向于对操作人员的认知过程进行建模, 基于认知模型对人为差错进行分类。CREAM (Cognitive Reliability and Error Analysis Method) 方法是其中的典型代表^[4]。CREAM 方法将认知过程分为四个模块: 观察-解释-计划-执行。该方法针对这四个模块, 分别给出了相关的差错类别。以解释模块为例, 对应的人为差错可以分为三类, 如表 1 所示。

表 1 解释模块对应的人为差错

Tab. 1 Human errors of explanation module

解释模块	11. 诊断错误——完全错误或部分错误
	12. 决策错误——应该决策而没有决策; 决策是错误的或者是不完整的
	13. 决策时间不合适——过早或者过晚

可以看出, 相对于前两种方法, CREAM 方法对人为差错的分类更为细致。它将认知模型引入到人为差错分类过程中, 分析不同的认知模块可能对应的人为差错, 符合人因可靠性分析研究的发展趋势,

也有助于分析人为差错的发生机理。

但是,需要指出的是,在 CREAM 方法中,确认了认知模块之后,直接就给出每个模块可能对应的人为差错模式。由于缺少两者之间的桥梁,因此,这一过程稍嫌突兀,而且由此得出的结果完全取决于方法设计者的个人经验和知识水平,在差错模式的全面性上说服力不够。如果能够找出一种基本的人为差错分类框架,用于指导不同认知模块下人为差错模式的分类,将是非常有意义的,如图 1 所示。

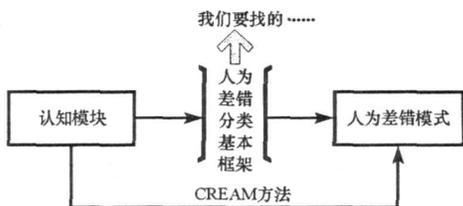


图 1 认知模块与人为差错模式
Fig. 1 The relationship between cognitive module and human error modes

2 人为差错分类基本框架

前已提及,执行差错/忽略尽管非常粗糙,但这种方法还是非常全面的。以此为基础进行细化可以设计一种比“执行差错/忽略”更为细致的分类框架,用于指导差错辨识过程,保证差错辨识结果的全面性。

执行差错可以从“空间”和“时间”两个方面进行细化。从空间的角度出发,如果执行动作没有达到任务要求,那么此时的动作就是一种执行差错;如果任务存在顺序要求,那么若顺序出现混乱,此时的差错也肯定属于执行差错。从时间的角度出发,如果任务存在时间要求,那么任务执行时刻过早或者过迟,都将属于执行差错。此外,如果任务是由一系列操作动作组成,若仅仅按照顺序执行了部分动作,此时的差错可以单独作为一类,即执行部分动作(非完全差错)。这样一来,可以得到如图 2 所示的人为差错基本分类框架。

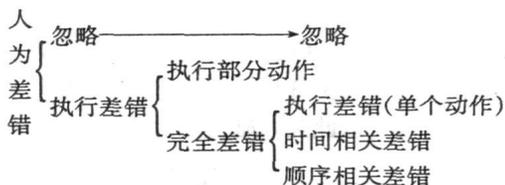


图 2 人为差错分类基本框架

Fig. 2 The framework for classifying human errors

由此可见,在“执行差错/忽略”的基础上,人为差错被细分为五类不同的基本模式,分别为忽略、执行部分动作、执行差错(单个动作)、时间相关差错和顺序相关差错。其中,忽略仍然表示“当任务要求人工干预时,没有采取任何动作”;执行部分动作指的是执行某个包含多个子任务的任务序列时,仅仅执行了其中的部分子任务;执行差错(单个动作)指的是执行单个动作时,所采取的动作是错误的;时间相关差错指的是当任务存在时间要求时,任务执行时刻过早或者过迟;顺序相关差错指的是执行某个包含多个子任务的任务序列时,子任务之间存在顺序要求,执行顺序出现混乱。

比如,某个任务需要相继执行动作 A 和动作 B,如果操作人员在执行动作 A 或 B 时出错,那么此时的差错就属于执行差错(单个动作);如果动作 A 需要在规定的时段内执行,那么在这个时段之外执行动作 A 就属于时间相关差错;如果操作人员先执行了动作 B 然后执行了动作 A,此时的差错则属于顺序差错。这些差错都属于执行差错这一大类。如果操作人员仅仅执行了动作 A 和动作 B 中的一个动作,那么此时属于执行部分动作差错。如果操作人员完全没有执行任何动作,那么此时则属于忽略差错。

这种统一的人为差错分类基本框架有以下两个方面的重要意义:

(1) 能够保证分析结果的全面性。如果没有统一分类框架的支持,仅仅依靠分析人员的经验来分析可能发生的人为差错模式,由于差错模式的多样性,因此难免会出现疏漏。而统一的分类框架由于涵盖了所有能够考虑到的差错类型,因此,在该框架的指导下进行系统和细致的分类,除非故意,否则要想遗漏差错类型也是比较困难的事情了。

(2) 能够保证分析结果的一致性。一致性包含两个方面的含义: 一是术语一致性; 二是结果一致性。人为差错存在多种表现形式, 在差错分析过程中如果没有统一框架的支持, 即使针对同一种差错, 不同的分析人员往往也会赋予不同的术语, 这将降低分析结果的可读性。此外, 没有统一框架的支持, 不同分析人员的结果可能将存在很大的差异, 导致差错分析结果的可信度低。

以上这些人为差错基本类别在辨识过程中起着指导作用, 结合具体任务时, 它们将以不同形式的差错描述语言体现出来。比如, 某个任务为“拉下阀门 A”, 那么此时忽略差错对应的差错描述语言则为“没有采取任何动作”; “拉下阀门 B”则属于执行差错(单个动作); “拉下阀门 A 过晚”则属于时间相关差错。由于这是一个单独动作, 因此, 不存在执行部分动作和顺序相关差错这两类差错模式。

3 基于认知模型的人为差错分类

图 2 所示的分类框架给出的基本差错类别可以作为人为差错类别的“原语”。基本的人为差错原语包括:

- 忽略
- 执行部分动作(非完全差错)
- 执行差错(单个动作)
- 时间相关差错
- 顺序相关差错

确定了人为差错原语之后, 接下来需要探讨如何将它们与认知模型相结合, 分析不同认知阶段可能出现的人为差错模式。

3.1 认知模型

将认知模型引入到人因可靠性分析中已经是这一领域的发展趋势。目前, 最为常用的认知模型是 Wickens 于 1992 年提出的“信息处理模型”, 如图 3 所示^[5]。在该模型中, 操作人员被看作是一个信息处理器, 首先从外界获取信息, 然后进行过滤分析, 以此为基础做出决策、制定计划, 最后依据计划执行相关动作。

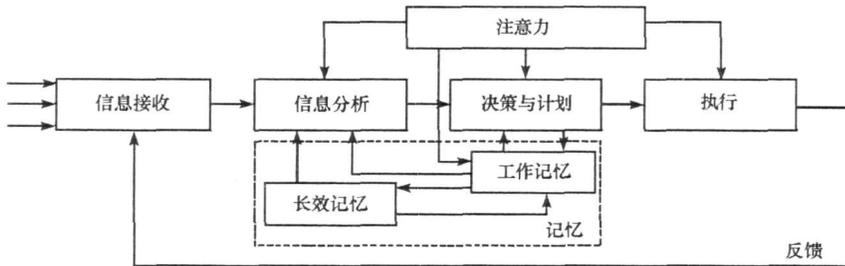


图 3 Wickens 的认知模型(信息处理模型)

Fig. 3 Wickens' cognitive model (The information processing model)

信息接收模块的作用在于被动或者主动地获取情景信息, 而信息分析模块则通过相关机理对信息进行过滤, 从而获取认为是“有用”的信息。接下来根据获取的信息判断系统状态, 做出相应的任务计划, 并提交执行。在这个过程中, 记忆资源能够存储相关信息, 注意力资源能够保证操作人员集中处理某个任务, 它们在认知过程中起辅助作用。在这四个不同的认知阶段中, 都有可能发生差错, 已经有很多研究人员针对不同的认知阶段分析可能发生的人为差错, 取得了一些有益的成果^[6-8]。

3.2 差错分类

套用上一节提出的差错分类框架, 借鉴已有的一些认知差错分类结果, 针对这四个认知阶段进行差错分类。具体分类结果如表 2 所示。

表 2 基于认知模型的人为差错分类结果

Tab. 2 Human error modes of various cognitive modules

认知阶段	差错大类	人为差错原语	差错描述
信息接收	信息接收差错	忽略	R1. 没有启动信息接收机理
		执行部分动作	R2. 接收到的信息不完整
		执行差错(单个动作)	R3. 没有找到应该接收的信息
			R4. 接收了错误的信息
		时间相关差错	R5. 没有及时接收信息
信息分析	信息分析差错	忽略	A1. 信息分析没有执行
		执行部分动作	A2. 遗漏了部分信息
		执行差错(单个动作)	A3. 信息分析出错(如过滤掉正确信息)
		时间相关差错	A4. 信息分析不及时
决策与计划	决策差错	忽略	D1. 决策与计划过程没有启动
		执行部分动作	D2. 决策与计划结果不完善
		执行差错(单个动作)	D3. 决策与计划完全错误
		时间相关差错	D4. 决策与计划过程启动过迟
		顺序相关差错	D5. 决策与计划过程在顺序逻辑上不正确
执行	执行差错	忽略	E1. 没有执行任何动作(应该干预的情况下)
		执行部分动作	E2. 执行的动作不够完整
		执行差错(单个动作)	E3. 执行的动作完全错误
		时间相关差错	E4. 执行的动作在时间上不合适
		顺序相关差错	E5. 执行的动作在顺序逻辑上不正确

由于存在人为差错原语的指导,相对于 CREAM 方法中给出的人为差错类别,上述差错分类的全面性更容易得到保证。得到以上人为差错类别之后,接下来就可以将它们应用到具体的任务中,确定在任务中可能出现的人为差错类型。

3.3 应用

在任务执行过程中,操作人员根据调用不同的认知模块来完成。需要指出的是,特定岗位的认知过程可能并不会完整地包含图 3 所示的所有认知模块。认知模块的调用取决于多种因素,如任务特性、工作能力和操作人员的角色要求等。因此,在分析人为差错时,

表 3 人为差错列表

Tab. 3 Human errors list

信息收集(信息接收模块)	信息上报(执行模块)
(1) 没有进行信息收集	(1) 没有上报信息
(2) 遗漏了一些信息	(2) 上报的信息不完整
(3) 没有及时收集信息	(3) 上报的信息与接收的信息不一致
	(4) 上报信息过迟

首先需要确定操作人员所调用的认知模块,然后同表 2 中给出的差错类别相对比,结合具体任务,得出具体的人为差错。

比如,存在某个岗位要求操作人员的任务就是收集所有信息并上报,此时操作人员无须决策直接收集信息上报即可。这种情况下,操作人员仅需调用信息接收模块和执行模块。此时,根据表 2 可得该岗位可能出现的人为差错模式,如表 3 所示。

4 总结

全面辨识与分析人为差错模式是 HRA 工作的基础。目前,人为差错辨识与分析基本上都依赖于分析人员的领域知识和主观判断,分析结果的全面性和一致性难以保证。针对这一点,本文在综合分析了现有的人为差错分类方法的基础上,提出了一种新的人为差错分类基本框架。这一框架包含五种基本人为差错模式,能够涵盖所有的人为差错模型。依据这一框架,结合认知模型,能够给出不同认知阶段可能出现的人为差错模式。针对具体任务时,首先分析操作人员所处的认知阶段,然后结合上述结果,分析可能发生的人为差错具体模式。由于存在统一的人为差错分类框架的指导,人为差错分析结果的全面性和一致性都能够得到保证。

6 结论

如何判断仿真模型之间是否能够组合, 在何种程度上能够组合, 是组合仿真方法中急需解决的关键问题之一。组合判定问题存在多个层次, 目前大多数基于组件的仿真都只能在语法层次上判断组件之间是否可以组合, 但仅有语法组合对于保证组合过程的正确性及合理性是不够的, 本文试图在模型语义层次上建立判定可组合性的方法, 主要工作集中在模型的语义描述、组合相容性、可替换性分析及其组合匹配规则。

本文以 Hoare 逻辑作为描述模型语义的方法, 并据此建立语义组合判定方法。由于 Hoare 逻辑只能描述模型单次计算即原子行为的语义约束, 而无法描述模型的行为序列(可能不终止)的语义性质, 因此本文研究范围限定在模型原子行为层次, 将模型在时间与空间上孤立起来分析组合过程的特征与约束, 空间上孤立是指不考虑两个模型与外部环境之间的交互作用, 时间上孤立是指只研究连续交互行为序列中原子行为所表现的动态语义之间的相容性及可替换性, 并且只考虑两个模型的组合, 没有涉及多个模型之间的组合问题。下一步的研究工作将逐步集中在模型行为序列的动态语义间的组合判定, 分布并发环境下模型间组合机制等问题。

参考文献:

- [1] Morse K L, Petty M D, Reynolds P F, et al. Findings and Recommendations from the 2003 Composable Mission Space Environments Workshop [C]//Proceedings of the Spring Simulation Interoperability Workshop, Arlington, VA, 2004.
- [2] 周东祥, 仲辉, 李群, 等. 复杂系统仿真的可组合问题研究综述[J]. 系统仿真学报, 2007, 19 (8).
- [3] Weisel E W, Petty M D, Mielke R R. A Survey of Engineering Approaches to Composability [C]//Proceedings of the Spring Simulation Interoperability Workshop. Arlington, VA, 2004.
- [4] Petty M D, Weisel E W, Mielke R R. A Formal Approach to Composability [C]//Proceedings of the Interservice/ Industry Training, Simulation and Education Conference Orlando, FL, 2003.
- [5] 陈意云. 形式语义学基础[M]. 合肥: 中国科技大学出版社, 1994.
- [6] Zeigler B P, Praehofer H, Kim T G. Theory of Modeling and Simulation: Integrating Discrete Event and Continuous Complex Dynamic Systems [M]. San Diego, CA: Academic Press, 2000.
- [7] Zaremski A M, WING J M. Specification Matching of Software Components [J]. ACM Transactions on Software Engineering and Methodology, 1997, 6 (4): 333- 369.

(上接第 77 页)

参考文献:

- [1] Kirwan B. A Guide to Practical Human Reliability Assessment [M]. Taylor & Francis, 1994.
- [2] Swain A D, Guttman H E. Handbook of Human Reliability Analysis with Emphasis on Nuclear Power Plant Applications (NUREG/CR- 1278) [R]. Washington, DC: US Nuclear Regulatory Commission, 1983.
- [3] Reason J, Human Error [M]. Cambridge University Press, New York, 1990.
- [4] Hollnagel E. Cognitive Reliability and Error Analysis Method (CREAM) [M]. Elsevier Science Ltd, 1998.
- [5] Wickens C D, Hollands J G. 工程心理学与人的作业[M]. 朱祖祥, 葛列众, 张智军, 等译. 上海: 华东师范大学出版社, 2003.
- [6] Shorrock S T, Errors of Perception in Air Traffic Control [J]. Safety Science, 2007, 45(8): 890- 904.
- [7] Shorrock S T, Errors of Memory in Air Traffic Control [J]. Safety Science, 2005, 43(8): 571- 588.
- [8] Zhang J J, Patel V L, Johnson T R, et al. A Cognitive Taxonomy of Medical Errors [J]. Journal of Biomedical Informatics, 2004, 37(2): 193- 204.