A Formal Approach to the Analysis of Human-Machine Interaction with Fuzzy Logic

Leyi Cui
Columbia University
New York, USA
lc3542@columbia.edu

Abstract

Human uses mental models when interacting with systems. Misalignment between a mental model and the system design, known as *mode confusions*, can lead to automation surprises. To better handle the vagueness of mental models through formalization, Fuzzy Mental Model Finite State Machines (FMMs), incorporating fuzzy logic, have been introduced. This work explores the potential of FMMs in formal analyses of human-machine interactions, proposing a set of formal behavior patterns of mode confusions and a tool for identifying mode confusions and unsafe interactions.

CCS Concepts: • **Human-centered computing** → **Systems and tools for interaction design**.

Keywords: Human Technology Interaction, Mental Model

ACM Reference Format:
Leyi Cui. 2024. A Formal Approach to the Analysis of Human-Machine Interaction with Fuzzy Logic. In *Companion Proceedings of the 2024 ACM SIGPLAN International Conference on Systems, Programming, Languages, and Applications: Software for Humanity (SPLASH Companion '24), October 20–25, 2024, Pasadena, CA, USA.* ACM, New York, NY, USA, 3 pages. https://doi.org/10.1145/3689491.3689969

1 Introduction

Human factors researchers have long argued that human use mental models to track the system's state and predict its behaviors [8]. Maintaining an accurate mental model is crucial for safe and effective human-machine interaction as misalignment between the model and the system, known as *mode confusions*, can result in *automation surprises*, where the system fails to act as expected [9]. Thus, avoiding mode confusions has become vital in the design of safety-critical systems and human-centered applications [4, 5, 7, 12, 14].

The executable nature of mental models has prompted some researchers to view them as rudimentary programs that can be represented by state machines [1, 10]. However, state machines with a deterministic execution semantics fall short of accurately modeling mental models because they do not capture the vagueness of human reasoning [8]. To overcome this limitation, researchers have introduced *Fuzzy Mental Model Finite State Machines (FMMs)* [2], a new formalism designed to precisely capture the vagueness and imprecision of human mental models. It builds upon state machines but incorporate fuzzy logic to handle vagueness. Fuzzy logic [15], developed to reflect the inherent imprecision of human thinking and language, allows for categories to have degrees of membership from 0 (not at all) to 1 (completely) rather than falling into discrete categories like true or false.

This work investigates the potential of FMMs in formal analyses to facilitate the verification of safe and robust human-machine interactions, explicitly accounting for the vagueness inherent in mental models. Contributions include:

1. A catalog of patterns of mode confusions.
2. A novel tool for analyzing human-machine interactions, identifying potentially unsafe behaviors and problematic vagueness within mental models.
3. A case study of applying the FMM Analysis Tool to user interfaces of contemporary mono-stable gearshifts.

2 Motivating Example

I examine a contemporary mono-stable gearshift system (Figure 1) designed to switch between four vehicle modes: parking (P), reverse (R), neutral (N), and drive (D).

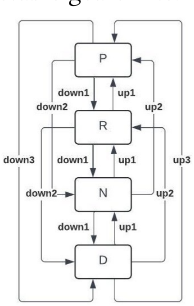

Figure 1. Gearshift system model.

Introduced in mid-2010s Fiat Chrysler vehicles, this design led to numerous accidents due to mode confusion, prompting a recall of over 1.1 million vehicles [6]. Unlike traditional gearshifts that remain in positions corresponding to specific gears, the mono-stable gearshift returns to its central position after each action. The design of this gearshift

Thank you to Dr. Eunsuk Kang at Carnegie Mellon University and Dr. Matthew L. Bolton at University of Virginia for advising me in this paper.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
SPLASH Companion '24, October 20–25, 2024, Pasadena, CA, USA
© 2024 Copyright held by the owner/author(s). Publication rights licensed to ACM.
ACM ISBN 979-8-4007-1214-2/24/10
https://doi.org/10.1145/3689491.3689969

can confuse driver due to the subtle feedback associated with each shift action and the absence of a fixed gear position. Specifically, the relatively indistinct feedback from each shift motion can fail to clearly indicate the number of shifts performed.

One type of accident that could occur in this scenario is when the driver is in neutral (N) and intends to reach parking (P) by shifting up once (up1). The driver mistakenly believe that by shifting up once the car should transits to parking (P), while the car is actually in reverse (R). This could cause unexpected backward movement, potentially leading to collisions, pedestrian danger, or loss of control.

3 FMM Analysis Tool

The analysis tool[1] takes two different models as inputs: (1) an FMM and (2) a labelled transition system (LTS) for the system model. It simulates state transitions of these two machines in parallel and identifies possible mode confusions.

3.1 Mental Model and System Specification

An LTS-based system model is a 4-tuple $S_2 = (\Sigma, Q, q_0, R)$, and an FMM is an 8-tuple: $S_1 = (\Sigma, Q, I, M, \vec{m}_0, \alpha, \phi, \delta)$ [2]. Σ denotes a finite set of input events that can cause a change in the system state, and Q represents a finite set of states the system could be in.

The user specifies the initial state $q_0 \in Q$ and the state transition function $R : Q \times \Sigma \rightarrow Q$ for the system model. Additionally, the user specifies the following for FMM:

- $\vec{m}_0 \in M$: a vector of initial state fuzzy set memberships.
- $\phi : Q \times \Sigma \rightarrow M$: a function that describes the fuzzification of state transitions by mapping current states Q and input events Σ to a vector of membership values for next states M.
- $\alpha : \Sigma \rightarrow I$: a function that describes how inputs are fuzzified by mapping crisp input events from Σ to a vector of input fuzzy set memberships from I.

For the motivating example, $\Sigma = \{up1, up2, up3, ..., down3\}$ and $Q = \{P, R, N, D\}$. The driver's belief about the modes that the vehicle could be in is represented by tuple $\vec{m} = (m_P, m_R, m_N, m_D) \in M$. It is assumed that initially, the driver is relatively certain that the vehicle is in the parking mode (i.e. $\vec{m}0 = (1.0, 0.01, 0.2, 0.01)$). As mentioned in Section 2, vagueness in the driver's mental model could come from confusion between the number of push-up motions, such vagueness can be represented in α. For example, when the driver pushes the shift up once (up1), they might confuse their motion as having done it twice (up2); this vagueness can be represented in α, where $\alpha(up1) = (0.9, 0.2, 0.0, 0.2, 0.0, 0.0)$.

3.2 Analysis

To find mode confusion errors, the tool randomly generates traces of a user-defined length, simulating the system's state transitions and the FMM' degrees of state membership.

[1] https://github.com/cmu-soda/FMMAnalysisTool

Table 1. An incomplete list of mode confusions.

Mode Confusion	Condition Over State Pair $(q, \vec{m})$
Dominant Error State	$proj_d(\vec{m}) = \{q'\} \wedge q \neq q'$
Nondeterministic Confusion	$\#proj_d(\vec{m}) \geq 1$
Vacuous Confusion	$\#proj_c(\vec{m}) = 0$

For input event $e1 \in \Sigma$, the tool calculates the degree of membership of the FMM as $\delta(\vec{m}, e1) = \vec{m}'$, where

$$\forall m'_x \in m' : m'_x = \vee_{m_y \in \vec{m}, i_{e2} \in \alpha(e1)} (m_y \wedge i_{e2} \wedge \phi(y, e2)_x) \quad (1)$$

where fuzzy **AND** is defined as $\wedge_{j=1}^{n} a_j = \prod_{j=1}^{n} a_j$ and fuzzy **OR** is defined as $\vee_{j=1}^{n} a_j = 1 - \prod_{j=1}^{n} (1 - a_j)$ [2, 13].

For example, given *up1* as $e1$, m'_P equals the **OR** ($\vee$) of all possible **AND** ($\wedge$) of current membership $m_y \in \vec{m}$ where $\vec{m} = (m_P, m_R, m_N, m_D)$, $i_{e2} \in \alpha(up1)$ where $\alpha(up1) = (0.9, 0.2, 0.0, 0.2, 0.0, 0.0)$, and $\phi(y, e2)_P$ given the $y, e2$ from m_y, i_{e2}. That is, given an input event *up1*, the degree of membership of state P in the next step $m'_P \in \vec{m'}$ can be calculated as above.

Given a pair of simulated states $(q, \vec{m})$, where $q \in Q$ is the current system state and $\vec{m} \in M$, the tool then analyzes these pairs of states for mode confusions as defined in Table 1, with $proj_c(\vec{m})$ defined as mapping $\vec{m}$ to a set of states whose membership value is greater or equal to c, and $proj_d(\vec{m})$ as mapping $\vec{m}$ to the state(s) with the highest membership. One type of mode confusion focused in this paper's discussion, *Dominant Error State*, refers to the scenario when the state that the human believes the system to be most possibly in does not match the system's actual state.

In the gear shifter example, the analysis tool identifies "N, up1 → P" as a common Dominant Error State error. This type of error occurs when drivers apply less force to shift gears than they realize or mistakenly believe they have already performed an action, leading to under-shifting [3, 11]. Similar errors are observed in all other Dominant Error State checks identified by the tool. The analysis tool detects the errors above with 1,000 simulations with a trace length of 10, requiring 25 seconds to complete. Similar errors can be identified by running fewer simulations or shorter traces, which significantly reduces the tool's execution time.

4 Conclusion and Future Work

I have presented a tool for analyzing human-machine interaction through FMMs, formal definitions of mode confusions, and a case study of gearshift system. Due to space limitations, additional contributions such as extended FMMs formalism, additional mode confusions and case studies are not included.

Future work includes developing elicitation techniques to construct FMMs from human-subject studies, creating a model checker for FMMs, exploring mental models in domains like aviation and medical devices, and identifying new design principles and automated repair methods using the FMM Analysis Tool.

References

[1] Matthew L Bolton, Ellen J Bass, and Radu I Siminiceanu. 2013. Using formal verification to evaluate human-automation interaction: A review. *IEEE Transactions on Systems, Man, and Cybernetics: Systems* 43, 3 (2013), 488–503.

[2] Matthew L Bolton, Elliot Biltekoff, and Kevin Byrne. 2022. Fuzzy mental model finite state machines: A mental modeling formalism for assessing mode confusion and human-machine "trust". In *2022 IEEE 3rd International Conference on Human-Machine Systems (ICHMS)*. IEEE, 1–4.

[3] KJ Harris and JK Jensen. 1964. John Deere power shift transmission. *SAE Transactions* (1964), 593–602.

[4] Erik Hollnagel and David D Woods. 2005. *Joint cognitive systems: Foundations of cognitive systems engineering*. CRC press.

[5] Adam Michael Houser. 2018. *Mental models for cybersecurity: A formal methods approach*. Ph. D. Dissertation. State University of New York at Buffalo.

[6] National Highway Traffic Safety Administration (NHTSA). 2016. *Investigation on Chrysler vehicle rollaway*. Technical Report EA 16-002. National Highway Traffic Safety Administration. https://www.autosafety.org/wp-content/uploads/2016/06/EA16-002-Chrysler-Rollaway-Closing-Resume.pdf

[7] Don Norman. 2013. *The design of everyday things: Revised and expanded edition*. Basic books.

[8] Donald A Norman. 2014. Some observations on mental models. In *Mental models*. Psychology Press, 15–22.

[9] Nadine B Sarter and David D Woods. 1995. How in the world did we ever get into that mode? Mode error and awareness in supervisory control. *Human factors* 37, 1 (1995), 5–19.

[10] Benjamin Weyers, Judy Bowen, Alan Dix, and Philippe Palanque. 2017. *The handbook of formal methods in human-computer interaction*. Springer.

[11] Jonathan C Wheals, C Crewe, M Ramsbottom, S Rook, and M Westby. 2002. Automated manual transmissions-a european survey and proposed quality shift metrics. *SAE Transactions* (2002), 1325–1343.

[12] Christopher D Wickens, William S Helton, Justin G Hollands, and Simon Banbury. 2021. *Engineering psychology and human performance*. Routledge.

[13] Mark J Wierman. 2010. An introduction to the mathematics of uncertainty. *Creighton University* (2010), 149–150.

[14] David D Woods and Erik Hollnagel. 2006. *Joint cognitive systems: Patterns in cognitive systems engineering*. CRC press.

[15] Lotfi Asker Zadeh, George J Klir, and Bo Yuan. 1996. *Fuzzy sets, fuzzy logic, and fuzzy systems: selected papers*. Vol. 6. World scientific.

Received 2024-07-08; accepted 2024-08-19