

IMPROVING HUMAN-MACHINE INTERACTION

A Thesis Submitted to the College of Graduate Studies and Research

In partial fulfillment of the requirements

For the Degree of Master of Science

In the Department of Mechanical Engineering

University of Saskatchewan

Saskatoon

By

Shrey Modi

August 2011

Permission to use:

In presenting this thesis in partial fulfillment of the requirements for a Postgraduate degree from the University of Saskatchewan, I agree that the Libraries of this University may make it freely available for inspection. I further agree that permission for copying of this thesis in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis work or, in their absence, by the Head of the Department or the Dean of the College in which my thesis work was done. It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any material in my thesis.

Requests for permission to copy or to make other use of material in this thesis in whole or part should be addressed to:

Head of the Department of Mechanical Engineering

University of Saskatchewan

Saskatoon, Saskatchewan (S7N5A9)

Abstract

This thesis studies human and machine interaction. For better interaction between humans and machines, this thesis aims to address three issues that remain unanswered in literature. Three objectives are proposed in this thesis to address the three issues, and the objectives are: (i) identification of the core capabilities of a Human Assistance System (HAS) and study of implementation strategy of the core capabilities; (ii) development of a framework for improving the accuracy of human mind state inference; (iii) study of the effect of representation of the machine's state (which is represented in a "natural" way) on the user's actions. By a natural way, it is meant a way that contains emotions known to be always present in humans (or human emotions in short).

The study includes theoretical development, experimentation, and prototype implementation. This thesis has concluded: (1) the core capabilities to be addressed in designing a HAS are transparency, communication, rationale, cognition and task-sharing and they can be implemented with the existing technologies including fuzzy logics, Petri Net and ACT-R (Adaptive Control of Thought-Rational); (2) expert opinion elicitation technique is a promising method to construct a more general framework for integrating various algorithms on human state inference; (3) there is a significant effect of the representation of the machine's state on the user's actions.

The main contributions of this thesis are: (1) provision of a case study for the proof-of-concept of HAS in the area of Computer Aided Design (CAD); (2) provision of an integrated framework for fatigue inference for improved accuracy, being readily generalized to inference of other mind states; (3) generation of a new knowledge

regarding the effect of the natural representation of a machine's states on the user's actions.

These contributions are significant in human-machine science and technology. The first contribution may lead to the development of a new generation CAD system in the near future. The second contribution provides a much powerful technology for human mind inference, which is a key capability in HAS, and the third contribution enriches the science of human-machine interaction and will give impact to the field of Artificial Intelligence (AI) as well. The application of the result of this thesis is rehabilitation, machine learning, etc.

Acknowledgements

I would like to express my heart-felt thanks to Chris, who in many ways is first a friend to his students and a mentor later. I am extremely grateful to him for his expert supervision and support he has given to me.

I would also like to express my gratitude to my committee members, Dr. Richard Burton and Dr. Daniel Chen, for their valuable suggestions during my committee meeting. I extend my thanks to all the members of the Advanced Engineering Design Laboratory for their priceless suggestions during various seminars.

I thank Tanu and Raj for the several brainstorming sessions in solving several issues related to this thesis. I am also thankful to my friends, Prathamesh, Nikhil, Arun, Udhaya, Sudha, Lily, Cindy, Forrest and Sougat for their support and encouragement. I wish the very best in life to all of you. I thank Kelley for being so helpful throughout my master's study.

I am grateful to my family members for their support and constant belief in me. I am also thankful to two wonderful people in my life, Nivedita and Pooja, for being there for me, whenever I needed them. Lastly, I thank the NSERC and the Department of Mechanical Engineering for the financial support.

Dedicated to my parents.

Publications related to this thesis

The following papers were published as a result of the research conducted during the course of this thesis study:

- 1) S. Modi, M. K. Tiwari, Y. Lin and W. J. Zhang, (2011), “On the architecture of human-centered CAD agent system”, *Computer Aided Design*, 43(2), pp. 170-179, Feb. 2011
- 2) S. Modi, Y. Lin, L. Cheng, G. Yang, L. Liu and W. J. Zhang (2011), “A socially inspired framework for human state inference using expert opinion integration”, Accepted for publication in *IEEE/ASME Transactions on Mechatronics*, Special Issue on Sensing Technologies for Biomechatronics, Oct. 2011

Table of contents

Abstract.....	II
Acknowledgements	IV
Publications related to this thesis	VI
Table of contents	VII
List of tables.....	X
List of figures.....	XI
Chapter 1: Introduction	1
1.1 Human-machine interaction	1
1.2 Motivation and research questions	1
1.3 Objectives	4
1.4 Organization of the thesis.....	5
Chapter 2: Background and Literature Review	7
2.1 Introduction	7
2.2 Human Assistance Systems (HAS)	7
2.3.1 HAS in Computer-Aided Design (CAD).....	9
2.3.1.1 Findings	13
2.4 Human state inference	14
2.5 Machine state representation	17
Chapter 3: Human Assistance System for Computer-Aided Design	20
3.1 Introduction	20
3.2 Tools used.....	21
3.1 ACT-R.....	21
3.2 Petri nets.....	21
3.3 Proposed framework.....	23
3.4 Example, results and discussion	36
3.5 Limitations of the proposed architecture	41
3.5 Conclusion.....	42

Chapter 4: An integrated framework for human state inference	44
4.1 Introduction	44
4.2 Proposed framework.....	45
4.2.1 Procedure I	45
4.2.2 Procedure II.....	46
4.2.3 Procedure III	47
4.2.4 Procedure IV	48
4.2.5 Information fusion	49
4.3 Implementation.....	49
4.3.1 Example	49
4.3.2 Experts used	51
4.3.2.1 Statistical experts (STAT)	52
4.3.2.2 Artificial Neural Network based experts (ANN).....	53
4.3.3.3 Static Bayesian Network based experts (SBN)	53
4.3.3.4 Takagi-Sugeno-Kang based experts (TSK).....	54
4.3.3.5 Performance of individual experts.....	54
4.3.3 Performance of the framework	55
4.3.4 Observations	57
4.4 Conclusion.....	59
Chapter 5: Effect of representation of machine’s state on user’s actions	61
5.1 Introduction	61
5.2 Simulated test-bed	61
5.2.1 Basic components of the interface	62
5.2.1.1 Component A.....	63
5.2.1.2 Component B	65
5.2.1.3 Component C	66
5.2.1.4 Component D.....	66
5.2.1.6 Component F	67
5.2.1.7 Component G.....	68
5.2.2 Versions of the interface	69
5.2.2.1 Version 1 – Traditional (T).....	69
5.2.2.2 Version 2 – Eliminated Candidate 1 (EC1).....	70
5.2.2.3 Version 3 – Eliminated Candidate 2 (EC2).....	71
5.2.2.4 Version 4 – Augmented Candidate 1 (AC1).....	72
5.2.2.5 Version 5 – Augmented Candidate 2 (AC2).....	73

5.2.3 Additional information about the game	73
5.2.3.1 Task.....	73
5.2.3.2 Measures	74
5.3 Experimental design.....	74
5.3.1 Participants.....	75
5.3.2 Experimental Procedure.....	75
5.4 Results	76
5.5 Conclusion.....	84
Chapter 6: Conclusion	85
6.1 Overview	85
6.2 Conclusions	85
6.3 Contributions	86
6.4 Limitations and future work	87
References.....	89
Appendix A: Certificate of ethics approval for the experiment	97
Appendix B: Data used for implementation of the framework for fatigue inference	98
Appendix C: Information fusion using OWA technique	102

List of tables

Table 3.1: Modified representation of the chunks of Moss et al. (2004).....	25
Table 3.2: Conversion of the fuzzy variable to the synthetic variable x.....	29
Table 3.3: Association rules of x with the fuzzy sets high, medium and low.	30
Table 3.4: Rule-table of the software system, which represents the “design rationale”...	31
Table 3.5: Pseudo Code.	35
Table 3.6: An example design chunk (adapted from Moss et al. (2004)).....	38
Table 3.7: A sample design chunk as represented in the model.	39
Table 3.8: An example of the elimination of loops by the CA.	39
Table 4.1: Mean and variance of the error in inference by individual expert for five runs	55
Table 4.2: Combination of experts.....	56
Table 4.3: Mean and variance of the error (%) in inference by the various combinations for three runs	57
Table 5.1: Messages displayed in Component F	67
Table 5.2: Details of Component G	68
Table 5.3: Details of the five versions of the interface.....	69
Table 5.4: Blocks with various treatments.....	75
Table 5.5: Paired differences between the pair EC1 and EC2.....	77
Table 5.6: Paired differences between the pair EC2 and AC2	80

List of figures

Figure 2.1: Generation of inference model from training data	14
Figure 2.2: Estimation of human state for a test case	15
Figure 3.1: A sample Petri net depicting places and transitions.	22
Figure 3.2: An example of the RB-PN implemented in the current work.	27
Figure 3.3: A sample table-design problem being solved and shown by use of the conceptualized Petri Net interface	28
Figure 3.4: The Multi-Agent Structure.	33
Figure 3.5: Flow chart of the architecture.	35
Figure 3.6: Cost of the transitions obtained in the iterations.	40
Figure 3.7: Cost of the designs obtained with prior transition knowledge.	41
Figure 3.8: Average cost of the solutions obtained with and without transition knowledge.	41
Figure 4.1: Framework of Procedure I. The problem is first viewed by the same category of experts from all perspectives.	46
Figure 4.2: Framework of Procedure II. The problem is first viewed by all experts from the same perspective (cue).	47
Figure 4.3: Framework of Procedure III. The opinions are fused together without any hierarchy	48
Figure 4.4: Framework of Procedure IV. The opinions from other procedures are fused together.	49
Figure 4.5: Fatigue inference model proposed by Yang et al. (2008)	50
Figure 5.1: Components of the interface of the test-bed.....	63
Figure 5.2: Locations of the fuel stations.....	64
Figure 5.3: Final display of the component A	65
Figure 5.4: Display of component B.....	66
Figure 5.5: Inactive and active states of the control button in Component C.....	66
Figure 5.6: Component D – low fuel icon	66
Figure 5.7: Component E – Fuel status of the vehicle.....	67

Figure 5.8: Version T	70
Figure 5.9: Version EC1	71
Figure 5.10: Version EC2	72
Figure 5.11: Version AC1	73
Figure 5.12: Average fuel level of the vehicle in EC1 and EC2.....	77
Figure 5.13: Average fuel level of the vehicle at refueling stations in EC1 and EC2	78
Figure 5.14: Total time spent refueling the vehicle during the task in EC2 and AC2.....	80
Figure 5.15: Number of times fuel was finished in EC2 and AC2	81
Figure 5.16: Preference of the interfaces amongst all interfaces	83
Figure 5.17: Preference of the interfaces in two clusters of polls.....	83

Chapter 1: Introduction

1.1 Human-machine interaction

Human-machine interaction (HMI) is a field of study of the interaction between humans (users) and machines (Sheridan and Parasuraman 2005). Research in this field leads towards better human-machine systems designed to perform tasks collaboratively with human beings. Modern assembly lines, airplane cockpit designs, everyday appliances, advanced driver-assist and in-vehicle systems, and software systems (e.g., computer-aided design software system, etc.) are all the result of the growing interest of both organizations and researchers in the field of HMI. It is an interdisciplinary field of human factors and ergonomics, computer science, engineering design, mechanical engineering, psychology and interface design.

1.2 Motivation and research questions

Recently, there are several issues raised by researchers across disciplines that need to be focused upon for future advances in HMI. One major issue is regarding the fundamental things that need to be considered by designers in the process of development of machines and other automated systems with which humans interact. For example, current computer-aided design (CAD) software systems for solving configuration design problems eliminate the humans from the design process instead of augmenting them. The process is becoming a totally automated process that essentially eliminates the human out of the loop. In this process human designers lose an opportunity of exhibiting their creativity. Another problem with the current HMI systems is that communication

between the human and the machine is very primitive and primarily is a monologue of information exchange. As Norman (2004) pointed out, two monologues do not make a dialogue. It is crucial, as a first step towards building better human-machine systems, to understand the core capabilities that are to be expected of machines in interaction with humans. Further, it is vital to critically assess several examples of human-machine systems from the viewpoint of these cardinal features. This exercise would yield a blueprint of basic functionalities that engineers and designers could use as a protocol to build good human-machine systems. Further there is no study in literature, which has focused upon design of a HAS from the perspective of these features. In this thesis, the domain of CAD system is studied to explore the design of a HAS. Therefore the first question in this thesis study is: **what are the basic cardinals to be addressed when building human-assistance systems for better HMI and how to implement such a system for CAD?**

Within the context of human-machine systems, there is a new paradigm for better interaction, which is to develop adaptive machines that can sense the user's state and act accordingly. For example, by means of sensors "planted" on the steering wheel of an intelligent vehicle, an on-board computer could recognize the emotional/physiological state of the human driver and take actions to put influence over the driver to change it properly (Lin et al. 2007). If it senses that the driver is in a drowsy state, it might possibly start to play the radio to keep the driver awake. If the car recognizes that the driver is very sad, it may play a driver's favorite song to cheer up him/her. This could lead towards a lower number of road accidents (Nass et al. 2005). Imagine a computer could feel our frustration and provide simple steps to do a task, which we wish to accomplish.

Research reports in literature seem to bring the future closer to reality by the day. However, a major issue remains that the state of the human cannot be accurately sensed. The main shortcoming is that if a machine takes action based on an incorrectly identified state of the human, it may lead to huge problems. There are several algorithms reported in literature to infer the emotional state of the human being. All these differ in the technique of inference and their power of accuracy. The problem is that each of these algorithms gives a so-called “opinion” about the state, which is only an estimate of the actual state. This issue calls for a method to increase the effectiveness in sensing the state of the human being. However, this method should not be just another algorithm to infer the state of the human with a goal to be more powerful in terms of accuracy. Instead, an alternative viable approach is to integrate the existing works to collectively yield a better-inferred estimate of the human’s state. This is because all of these algorithms are candidate estimates of the human’s state and an approach is needed to integrate these opinions from varied sources. Hence, the second question raised in this thesis is: **Is there a way to integrate various techniques of inference of the human’s mental or emotional state for better accuracy and sound principle of inference?**

Recently a new paradigm is being raised in HMI towards the goal of creating better machines, which is to realize the importance of representation of machine state in a natural way that is easier to understand by humans. This has far-reaching implications and great potential to change the fundamental way that humans interact with machines. What if machines could have and express emotions? What if machines could represent themselves in a natural way? These questions were raised by Norman in his book, ‘The design of future things’, where he brilliantly summarized why emotions are necessary for

machines (Norman 2007). He also proposes that if the machine could communicate about its state in a natural manner, it may have an effect on the human interacting with the machine. It might bring a change in the actions of the user of the machine and may affect the human's long-term behavior. However, there is no proof-of-concept study in support of his proposal. The third question raised in this thesis is: **would representation of machine's state in a 'natural' way significantly affect the user's actions in doing a task?**

In general, the contemporary literature in HMI has not provided sufficient knowledge to answer the aforementioned three questions; a detailed review of literature is provided in the next chapter. The motivation of this thesis is to generate knowledge to answer these questions and is to advance technologies for better HMI.

1.3 Objectives

There are three specific objectives defined for this thesis study:

Objective 1: Explore the core capabilities in human-machine systems for good interactions between humans and machines by means of a case study about HAS and design a prototype HAS for CAD from the perspective of these core capabilities. The work along this objective is to provide answer to the first question as mentioned before. It is noted that the machine system is an intelligent CAD software system. However, the outcome can be easily adapted to any so-called intelligent hardware machine system.

Objective 2: Develop a framework that allows the integration of various algorithms/techniques for human state inference. The work along this objective is to provide answer to the second question as mentioned before. The idea behind the work on

this objective is that hybridization and integration of various types of algorithms on multiple cues should result in a more accurate result than a single algorithm. This idea is generalized from the work of Yang et al. (2008), Lin (2010) and Zhang et al. (2011).

Objective 3: Build a prototype simulated test-bed in which the machine can communicate with the human about the machine's state in a natural way; design and conduct experiments to generate knowledge about the effect of such a natural expression of a machine's state on user's actions. The work along this objective is to provide answer to the third question as mentioned before.

1.4 Organization of the thesis

This thesis is comprised of six chapters. The subsequent chapters are organized as follows:

Chapter 2: This chapter introduces the basic knowledge about Human-Machine Interaction (HMI) and other related fields. The three issues addressed in this thesis are extensively commented upon to establish the context of the proposed research further.

Chapter 3: This chapter builds upon the first issue and hence achieves the first objective of this study. The core capabilities of a general-purpose HAS are identified by means of a case study in the area of computer-aided design.

Chapter 4: This chapter presents the work related to the second objective. It introduces the architecture of the proposed framework to integrate various algorithms/techniques for human state inference. Fatigue is taken as an example to illustrate the effectiveness of the proposed architecture. The results obtained are also discussed for their further implications.

Chapter 5: This chapter relates to the achievement of the third objective. A detailed description is provided about the prototype developed for the study. The experimental procedure and results are also discussed to understand the specific effects.

Chapter 6: The final chapter presents the conclusions from this work. Contributions, limitations and future research directions are also discussed.

Chapter 2: Background and Literature Review

2.1 Introduction

This chapter provides literature pertinent to the objectives as defined in Chapter 1 of this thesis. Section 2.2 discusses the literature about HMI. The issues related to human-in-the-loop are examined in Section 2.3. Reported studies in the area of human state inference are studied in Section 2.4. Various approaches in representation of machine state with their benefits are reasoned in Section 2.5. At the end, the research objectives defined in Chapter 1 are revisited to further justify the need of the proposed research and its nature.

2.2 Human Assistance Systems (HAS)

Human-machine interaction (HMI) is a field of research that focuses on the study of interaction between humans and machines. Recently, an emphasis is being given in research literature to make designs of human-machine systems that keep the human ‘in-the-loop’ in HMI (Ruff et al. 2002). This means that the human should be also an active participant in the control and management of tasks in a work environment. This emphasis may be seen as being further rooted in two paradigms in systems science and engineering, namely (a) machine’s intelligence can never surpass human’s intelligence; and (b) machine may go wrong in a similar manner as humans can make mistakes. Following paradigm (a), keeping the human ‘in-the-loop’ will literally give an unlimited possibility of intelligence of the underlying human-machine system. Paradigm (b) reminds us of the ultimate duty of human beings – in particular when the machine goes wrong, the human has to take over the machine’s job to keep life moving on.

The school of thought of human ‘in-the-loop’ in HMI leads us towards the study of Human Assistance Systems (HAS). These systems are conceptualized to be assisting humans in their tasks by offering situation based help (Cai 2011). The notion is that these systems augment humans in a task rather than eliminating humans.

Currently, there is no specific set of capabilities outlined in literature that designers can use to build HAS with better HMI in mind. Following is a set of five capabilities which seem crucial for a HAS to possess:

- 1) Transparency
- 2) Communication
- 3) Rationale
- 4) Task-sharing
- 5) Cognition

The aforementioned list is far from being a complete and exhaustive list of cardinal points that need to be considered by designers when building these useful systems. However, they seem to be the core capabilities that are desired in any system built for interaction with humans. It is essential that several systems be analyzed from the viewpoint of the above to create a final blueprint of the desired capabilities. In the scope of this thesis, a case study of human-machine system is considered and it is scrutinized from the perspectives of the above capabilities. The case is in the domain of CAD system in which increasingly, agent-based systems are being created for assisting designers. Based on the critical analysis, shortcomings in the existing literature are identified in the subsequent sub-sections of this chapter. Possible frameworks for addressing the shortcomings are presented as solutions in Chapter-3.

2.3.1 HAS in Computer-Aided Design (CAD)

This case is about humans and software systems working in cooperation to generate conceptual design solutions to a given engineering design problem, configuration design in this case. It is hypothesized that a HAS for providing assistance in CAD possess capabilities such as communication, transparency etc. This falls under the contemporary definition of use of computational agents, which are sophisticated programs that possess attributes such as social behavior, reactivity, etc. and can solve tasks on behalf of their users (Franklin and Graesser 1997). Increasingly, complex tasks require multiple agents that can work together. A multi-agent system (MAS) is a collection of multiple agents, possessing diverse capabilities, which work interactively, distributing various resources amongst them, to achieve a common goal (Lang and O’Gardy 2001).

In the current context, the capabilities the HAS needs to have can be interpreted as follows:

- 1) Task sharing: This refers to the sharing of tasks/sub-tasks in the design process among the participants;
- 2) Transparency: The capability of the computer agent to represent its actions to the designer so that the designer can be ‘in-the-loop’;
- 3) Rationale: The capability of the computer agent to have appropriate rationale for every step it executes in the design synthesis in particular;
- 4) Communication: The capability of the agent to communicate with and receive feedback from the human designer through a human-agent interface;

- 5) Cognition: The capability of the agent to learn from the past design experience.

Agent-based design models are reported in the literature. A-Design (Campbell 2000, Campbell et al. 1999, Campbell et al. 2000) is an iterative, multi-agent-based design for solving engineering design problems. It consists of four classes of interacting agents – The Configuration agents (C-Agents) construct candidate designs using a library of embodiments depending on the input and output constraints specified by the human designer; the Instantiation agents (I-Agents) take the configuration designs from C-agents and instantiate the parameters in the system with values obtained from a catalogue of components; the Manager agents (M-agents) take the design population and produce feedback that controls how other agents in the system operate; the Fragmentation agents (F-agents) take out one or more components from the good designs and these fragmented designs are reconstructed and become part of the next iteration’s design population.

A-Design system has a primitive level of communication capabilities: it consists of two monologues between the human designer and the agent system: the designer issues the command and agent displays the result. Also, the rationale built in the system is to heuristically search and connect inter-connectable components to complete the design. However, this can be considered as a kind of operating logic that is the behavior of the system. The “actual” rationale should be the rationale behind choice of the components, which in this case is a heuristic search. This serves very little value, as the rationale of this kind is not understandable to the human designer.

Establishing useful and successful communications between the computer software and the human designer has also been demonstrated in literature in another

agent-based design system proposed by Campbell et al. (2003). In this system, human designers can interrupt the design synthesis and initiate a small dialog with the Manager-agent if he/she wishes to, by which he can adjust the rating system of the design process. The dialog consists of a task for the human designers to rate three design variants on a scale of 1 to 10 provided that the best solution in a particular iteration is considered to have a rating of 5. It is worthwhile to note that in the agent-based design in Campbell et al. (2003), there is a notion of human designers and agents, cooperating with each other to produce the designs. In other words, there is a level of automation built into the system such that the system is not fully automated. The sharing of tasks here is that the human designer uses his experience, to guide the agent to produce good designs, by teaching the agent to rate the design variants. Another notable improvement of this system is the increased level of communication capabilities. However, the capabilities of transparency and cognition are still lacking in this system.

Moss et al. (2004) proposed a software system in which the agents possess the ability to learn from design experience. This system shows to be very effective in deploying the experience by transferring the knowledge learnt in the previous iterations. Design processes for three electromechanical problems were demonstrated with this system. The ability of cognition is given to the agents by using the ACT-R (Adaptive Control of Thought-Rational) architecture (Anderson and Lebiere 1998). Although this system is successful in imparting cognitive abilities in terms of representing design components in the declarative module or declarative memory, the procedural memory remains a black box. The design synthesis process is highly event-driven; the type of component to be chosen in later stages of design depends on the component chosen in the

current design stage as the component functional parameters should match. However, the model of Moss et al. (2004) lacks a proper formalism to represent this complex event-driven behavior of the design system. In other words, the software system lacks the capability to make its design procedure transparent to the human designer. Also, at each stage, selection made between competing solutions depends on several parameters. For example, a structural frame with built-in fixtures for an air conditioner might eventually yield a low cost design than to buy the frame and fixtures separately. Hence at the design stage of choosing the frame, the human designer's past experience at the stage comes into play. Furthermore, the human designer does consider the cost of individual components while deciding different solution components, and this means that his/her decision is also influenced by his/her experience of how well the component would minimize the entire design cost. This issue is to do with logic or the rationale behind the choice of the components.

Applications of evolutionary algorithms for solving engineering design problems are also seen in the literature (Tay et al. 1998, Brown and Hwang 1993, Carlson-Skalak et al. 1998). However, evolutionary algorithm based approaches have three major disadvantages: 1) They fail to incorporate the human cognitive process of problem solving (i.e., capture the human activity in design); 2) They keep the human operator in the dark and act as a black box thus lacking transparency in representing the design synthesis; 3) They lack a good formalism to explain the rationale of their problem solving approach. In other words, the heuristic algorithmic approach to software design does not follow the human-centered design principle and would not yield a good HAS for design.

Researchers have also tried Object-Oriented approach for solving engineering design problems (Kusiak et al. 1991, Lang and O’Gardy 2001). However, modeling design knowledge in an object fashion is similar to the declarative module of the ACT-R architecture – the objects in the object oriented programming paradigm are the same as the chunks in ACT-R. The added benefit that ACT-R provides is to use this representation to solve problems in terms of procedural knowledge.

2.3.1.1 Findings

From the foregoing literature survey of the state-of-the-art in HAS in CAD software systems, the following findings can be made:

- 1) Multi-agent systems are the current trend in HAS for CAD as they have the best mimicking of the contemporary design practice – i.e., a team of designers collaborating with each other.
- 2) Researchers have explored the capabilities of communication, cognition and task sharing in HAS for CAD.
- 3) The capabilities of transparency and rationale have yet to be explored in a more pragmatic way in this domain of HAS.

The capabilities of cognition, transparency and rationale have unique meaning to the domain of engineering design. These have yet to be incorporated in CAD agent systems. In Chapter-3, a framework as well as the strategy to implement it is proposed to address the above limitation.

2.4 Human state inference

Human state inference has several important applications (Goleman 1995). Human state inference is generally done from physiological and contextual cues (or signals), which are believed to be indicators of the state. The inference system is a mapping of the cues to the state. Since there are no first principles in literature about this relationship, researchers take support of empirical learning to construct such a mapping. In this process, historic data, more commonly known as “training data”, about the cues and the human state is vital to be collected. As shown in Figure-2.1, training data is used to generate the inference model by use of machine learning techniques such as Artificial Neural Networks (ANN), Bayesian Networks (BN), etc. This inference model can be subsequently used to determine the state of a given test case as shown in Figure-2.2.

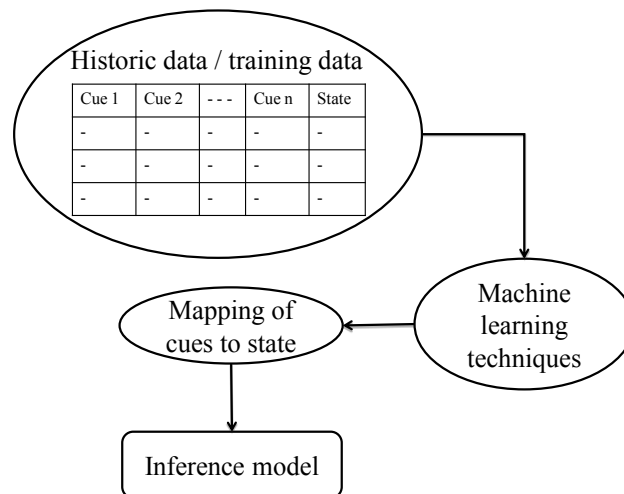


Figure 2.1: Generation of inference model from training data

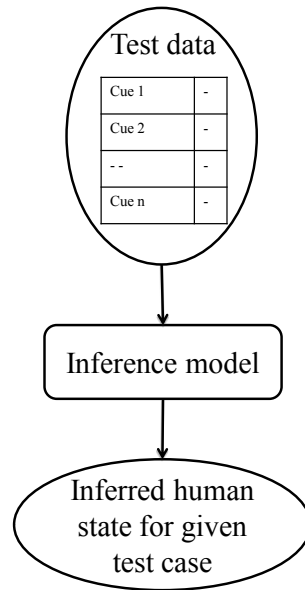


Figure 2.2: Estimation of human state for a test case

Picard et al. (2001) proposed a hybrid recognition algorithm to infer the emotional state of a person. Their mapping of physiological cues (such as muscle activity, skin conductance, heart rate variation and respiration rate) to eight emotional states achieved 81% accuracy. Nasoz et al. (2003) also reported an emotion inference system using physiological cues namely heart rate, skin temperature and skin conductance. In their work, they elicited emotions in the participants by use of various scenarios in different movie clips, which the subjects watched. Hence, their system in effect maps the cues to various scenarios. Three algorithms were used in their method to determine the weights: k-Nearest Neighbor (KNN), discriminate function analysis (DFA) and Marquardt back propagation (MBP). In their later work (Nasoz et al. 2004), the three said algorithms were combined with feature extraction technique to result in an overall improvement of accuracy of inference.

Mandryk and Atkins (2007) proposed a fuzzy inference system to infer the emotions of users in a gaming scenario. The cues used are galvanic skin response (GSR), heart rate (HR) and electromyography (EMG). Two fuzzy rule engines: First is to determine the arousal and valence of the player's experience from the cues; second is to determine the emotion from the arousal and valence obtained from the first set of fuzzy rules. The if-then rules were constructed using knowledge from a previous study of emotion (Russell et al., 1989). Mental workload was inferred and controlled by Wu et al. (2008) using the Queuing Network-Model Human Processor (QNMHP) technique (Liu et al. 2006). Human cognitive state has also been modeled using the Adaptive Control of Thought-Rational (ACT-R) (Anderson and Lebiere 1998).

Human fatigue state determination has also gained interest among the researchers. He and Zhao (1993) inferred the Fatigue level of human drivers by using fuzzy set theory. Hamonda and Saccomanno (1995) proposed a neural network model for fatigue detection for truck drivers. Jung et al. (1997) proposed a fuzzy-neural network based inference system to detect drowsiness of drivers using the signals from EEG. Horng et al. (2004) report a fatigue detection algorithm using the eye tracking technique. Yang et al. (2008) used a fuzzy neural network technique namely Takagi-Sugeno-Kang (TSK) to infer the fatigue state of the human drivers. Fatigue inference has also been achieved using Static Bayesian Networks (Ji et al. 2004) and Dynamic Bayesian Networks (Yang et al. 2010, Ji et al. 2006, Li and Ji 2005).

As evident from the aforementioned discussion, there are several algorithms (or techniques) reported in literature to infer the human state. Each of these techniques can be considered to be an expert in making his/her decision on the state. Each expert has

good and bad points. The problem is that, each of these experts gives a so-called ‘opinion’ about the human state. Each of these opinions is only an estimate of the actual state of the human. An integrated framework is necessary to be developed because none of the experts is powerful enough to infer the state for every situation (Gray 2007, Lin 2010). One of the driving forces to pursue this integrated framework is that if the machine takes action based on an incorrectly identified state of the human, it may lead to much bigger problems which is due to the poor human-machine interaction. An integrated framework might be able to advance the accuracy of inference.

2.5 Machine state representation

Representation of a machine’s state in a natural way has gained increased interest among researchers in the field of HMI and artificial intelligence. It is believed that this way will lead towards better communication between humans and machines. Bartneck (2001) developed expressions for a machine to display human-like emotions using facial expressions of an embodiment of a human face. Saerbeck et al. (2010) developed a robotic cat (iCat) that displays emotional expressions for use in tutoring applications. Kismet is a robot developed at MIT that uses movements of lips, ears, jaw etc. to display emotion (Breazeal 2002). Saldien et al. (2010) developed a robot that appears like an elephant in which the facial expressions of a cartoon character ‘Probo’ are brought to life in the context of the non-verbal way for representation of emotion. Suzuki et al. (1998) presented a mobile robot, which is hypothesized to feel certain emotions based on its interaction with the human by means of gestures. In particular, based on the gestures of the human, the emotional state of the robot changes, which further causes the robot change its movement in the environment.

Saerbeck and Bartneck (2010) proposed to use motion features, such as acceleration and curvature, of a robot to express emotions. The main focus of their work was to determine how humans perceive a robot's motion as emotion. Participants were asked by questions relating to which emotional label they would use to characterize the motion of the robot. They concluded that when the robot was moving with high acceleration, humans perceived it as an increase in the robot's arousal.

Use of emotions for machines has also applications in machine learning. Gadando and Hallam (1998) attempted to make a robot that takes decisions based on the emotion felt by the robot. Lee-Johnson and Carnegie (2010) presented a system in which robot is hypothesized to feel certain emotions. Using these emotions, an algorithm was developed for path planning of the robot.

Potkonjak et al. (2002a, 2002b) took a task of hand writing as an example and demonstrated the emotion of "fatigue" using a novel approach to link increasing motor temperatures to increasing fatigue. The temperature increases as the machine does repeated work, which in turn increases the fatigue. Their idea fits well with the definition of fatigue in the context where fatigue arises from repeated loading and unloading. The behavior of the robotic hand was also modeled such that it can change its methodology to achieve its goal (i.e., handwriting task) in elevated states of fatigue. In such scenarios, the performance of the robotic hand decreases, which is visible in the output of the task. For example, the writing is not very smooth when the hand is in high fatigue state. This is also an innovative way for the robot to communicate about its state to the human.

Nass et al. (2005) presented one of the first works on how machine's representation of its emotional state can improve the safety of the passenger. They proved

that pairing occurs between the driver's emotion and the car voice's emotion and that the latter can affect the former. They further studied the effect on the driver's emotion on the driver's performance in the task.

A natural way to represent of a machine's state would mean that there are certain labels that describe human emotions and they are linked to various states of the machines. It would be very easy for humans to relate to these labels. For example, if the engine temperature is high of a car, it can be called as the car being 'tired'. When the car is low on fuel, it can be represented as the car being 'thirsty'. Norman (2004) argued that such a representation would help to evolve the interactions between humans and machines. Humans might consider machines as a friend and take good care (i.e., maintenance) of the machines, when the humans understand the correspondence of a machine's state to a machine's natural expression (adapted from Norman 2007).

In summary, the above literature survey of machine state representation highlights an explored question: is there any significant effect on the user's actions in a task when the machine's state is represented in a natural manner. It is noted that this question has not been answered by the current literature. Nass et al. (2005) provided a partial 'yes' answer to this question; yet their work has not considered the machine's state. In fact, a hired actor generated the voice that was given out by the car (Nass et al. 2005), and there is no linkage between the voice and the machine's state. This thesis aims to answer this question, and in particular the answer is presented in Chapter-5.

Chapter 3: Human Assistance System for Computer-Aided Design

3.1 Introduction

In Section 2.3 of this thesis, a case study of a HAS for CAD system to support human designers in solving configuration design problems was analyzed. The major finding obtained is that the capabilities of transparency, rationale and cognition have not been studied in the context of CAD software. This chapter proposes a framework that addresses this issue. The framework uses a modified representation of the ACT-R chunks as originally used by Moss et al. (2004) to acquire the capability of cognition. The other capabilities of transparency and logic are added to the system by use of Petri nets and Fuzzy logic respectively. These capabilities have never been demonstrated in the literature of the HAS for CAD systems. The effectiveness of the proposed architecture is demonstrated by solving a hypothetical configuration design problem along with its implementation. This implementation also gives great confidence for a more comprehensive implementation of the proposed architecture for more complex design problems.

In Section-3.2, a brief introduction about the Petri nets and ACT-R is presented. The proposed framework is discussed in Section-3.3. The results of implementation of the framework to solve a hypothetical test problem are discussed in Section-3.4. In Section-3.5, a conclusion is presented towards achievement of the first objective defined in Chapter-1 of this thesis.

3.2 Tools used

3.1 ACT-R

ACT-R is a cognitive architecture to understand and simulate human cognition (Anderson and Lebiere 1998, Anderson et al. 1995, Lebiere and Anderson 1993). It consists of two types of memory modules namely, 'Declarative Memory' and 'Procedural Memory'. Knowledge is represented as chunks (or buffers) that collectively constitute the declarative memory. Procedural memory consists of production rules representing the knowledge of how humans do things (Anderson et al. 1995). It is essentially the most vital module of ACT-R, as it possesses the knowledge of performing the tasks (Anderson et al. 1997). For example, in the case of a calculation task, procedural memory is the generalised knowledge of the process of performing addition tasks making use of addition chunks (declarative memory) – the effectiveness of ACT-R depends more on the effectiveness of the procedural module, and hence it should be carefully modelled. For further information on the ACT-R architecture, the reader is directed to the bibliography maintained by the ACT-R group at CMU (<http://act-r.psy.cmu.edu/>).

3.2 Petri nets

Petri nets (PNs) provide a proper tool to explicitly represent process knowledge, and Fuzzy Logic is best suited for explanation of rationale of decision making in the case of competing solutions. The PN is a very promising tool to represent concurrent or parallel systems and can clearly capture the system constraints better than other models (Murata 1989). The PN may be viewed as a particular kind of directed graph (Figure-3.1) comprising of four kinds of primitive objects - places, transitions, directed arcs

connecting the places to the transitions and the transitions to the places and tokens. A detailed tutorial of PNs can be found in Zurawski and Zhou (1994).

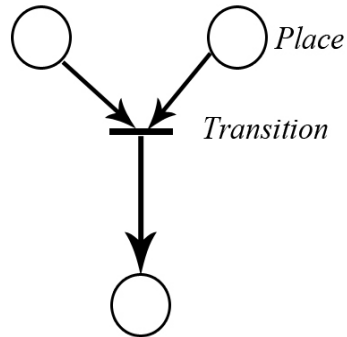


Figure 3.1: A sample Petri net depicting places and transitions.

A Petri net (PN) is represented in tuples, an example of which is illustrated below.

$$\text{PN} = (\text{P}, \text{T}, \text{I}, \text{O}, \text{M}) \quad (3-1)$$

where,

P = Set of Places,

T = Set of Transitions,

I = Input Function,

O = Output Function, and

M = Marking of the Petri Net.

The example shown above consists of 5 tuples. The number of tuples may be increased to incorporate additional information. Marking denotes the state of the PN by associating places with tokens. A transition can be enabled only when the number of

tokens at all the places is greater than the weight of the corresponding connecting arcs. However, this may not be the sole criteria of enabling a transition but it is the constraint for a particular transition to qualify for enabling. More information on Petri nets can be found in the seminal paper by Murata (1989).

3.3 Proposed framework

From a conceptual standpoint, it is proposed to use fuzzy rule based Petri nets for design rationale and transparency, and ACT-R model of Moss et al. (2004) for cognition. However, to discuss the architecture in detail, a context (or function) of the CAD system is necessary.

The proposed architecture describes a CAD agent system with the following functionality: given a set of input and output specifications of design requirements, the CAD system returns design solutions with a minimum cost. Hence the design cost is the objective function, which is to be optimized by the CAD agent system. It is to be noted that the scope of this objective function is limited to the parameters such as weight, cost, or cumulative error, in which the cumulative sum of the specified parameters of the components in the design solutions represents the parameters of the entire design itself. For example, the weight of a design solution or product is same as the sum of the weights of the individual components in the product. The proposed architecture is based on the aforementioned context (or function) of the CAD system. However, it is noted that the architecture can be modified appropriately to suit to other contexts.

In the proposed architecture, the human designer studies the problem and identifies the input and output constraints (domain and interface) required in the design.

He/she inputs these constraints into the CAD agent system and specifies an objective to which the function is to be optimized (minimization of the objective function). After the human designer inputs the required specifications, the agents take charge and search for designs iteratively until the number of iterations exhausts or the design cost reaches to an acceptable level – this criterion being specified by the human designer.

The remainder of this section first describes the modeling of the context of the problem using ACT-R and rule-based Petri Nets. Second, a short note is presented on the multi-agent architecture illustrating the interactions between the agents and their collective problem solving technique. Finally, the aspect of learning from knowledge transfer is discussed, and a pseudo-code of the entire architecture is presented.

The chunk representation, which constitutes the declarative part of ACT-R, in the context of engineering design problems, has been illustrated in the work of Moss et al. (2004). In the architecture proposed in this thesis, the cost attribute of the chunk is added to the representation presented in (Moss et al. 2004) as shown in Table-3.1, and this representation is used as the design chunk in the proposed model. This extension is necessary as the cost of the chunk is used in the rule based Petri net to take decisions which build the design rationale of the software to minimize the objective function, as discussed in the subsequent paragraphs.

The domain and interface attributes collectively represent the Functional Parameters (FPs) of the chunk. Two chunks can connect with each other only if the output FPs of the former chunk matches to the input FPs of the latter – e.g., in the case of a desktop computer, a monitor can only connect to the CPU via the socket provided in the CPU for monitor cable; this joint is the match of the output and input FPs of the monitor

and the CPU, respectively. A chunk consists of several components (Table-3.1) and the cost of the chunk is the cumulative sum of the cost of these individual components.

Design-Chunk C		
Isa	:	Design-chunk
Input-domain	:	Translation/rotation etc.
Input-Interface	:	Bolt/Shaft-hole etc.
Output-domain	:	Translation/rotation etc.
Output-Interface	:	Bolt/Shaft-hole etc.
Components	:	C1, C2 etc.
Connectivity	:	Port-1 of C1 is connected to Port-1 of C2.
Goodness	:	Cost of the chunk

Table 3.1: Modified representation of the chunks of Moss et al. (2004).

The computer agents start building the solution by selecting a chunk whose input FPs match with the input FPs specified by the human designer. At each stage of the design, including the initialization of a design, there might be several chunks that have the input FPs similar to the output FPs of the current stage of the design (and for the initialization chunk the input FPs need to be similar to the input FPs specified by the human) and decisions have to be made to select one of these chunks to proceed towards the completion of the design. In other words, the choice system is parallel and concurrent. Ideally, the analysis process has to be able to learn to build and take smarter decisions and this requires the decision evaluation system to be adaptive. A rule-based Petri net is chosen for such system in the current model because it offers a logical reasoning about the choice of decisions made and the whole analysis process is clear to the human

designer who is in control of the model; it offers a grey box if not a white box understanding and transparency. This rule based Petri net (RB-PN) accounts for the procedural knowledge (or the production system) of ACT-R. The RB-PN proposed in an 8-tuple PN as follows:

$$\text{RB-PN} = (\text{P}_{\text{st}}, \text{P}_{\text{so}}, \text{T}, \text{I}, \text{O}, \text{D}, \text{W}, \text{M}) \quad (3-2)$$

where,

P_{st} = Set of State Places (current state of the design),

P_{so} = Set of Source Places (chunks that can be added to the current state),

T = Set of Transitions (events which advance the design stage by firing),

I = Input function,

O = Output function,

D = Set of Propositions,

W = Preferences of the transitions, and

M = Marking of the Petri Net.

A primitive example is illustrated in Figure-3.2. The alphabets ‘A’, ‘X’, ‘F’, etc. in Figure-2 represent various chunks available in the knowledge pool (collection of design chunks) which also denote the places of the RB-PN. The places ‘A’, ‘A-X’, ‘A-F’, ‘A-X-F-M’, and ‘A-X-C-M’ denote the state places (P_{st}) and the places ‘X’, ‘F’, ‘P-M’, ‘C-M’ denote the source places (P_{so}). The transitions ‘ t_1 ’, ‘ t_2 ’, ‘ t_3 ’ and ‘ t_4 ’ represent the events which advance the design. The problem is to construct a solution which has the input and output FPs of the chunks ‘A’ and ‘M’, respectively. Such a design can be

obtained by either firing transitions of the set (t_1, t_3) or the set (t_2, t_4) , respectively, and there might be many more similar sets to construct the design.

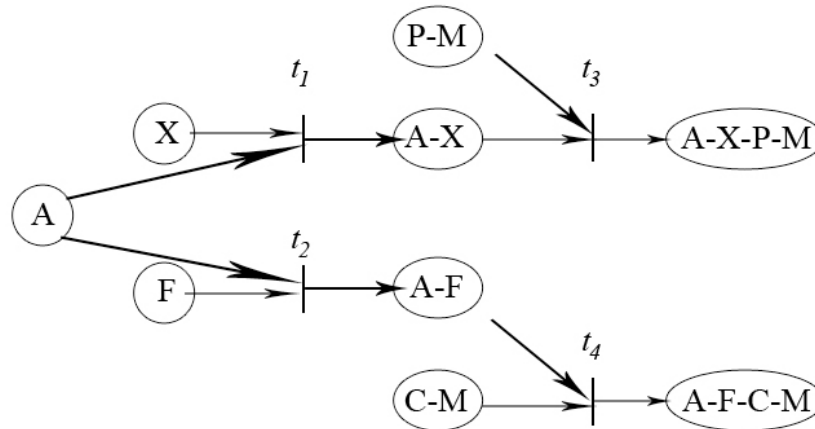


Figure 3.2: An example of the RB-PN implemented in the current work.

It is worthwhile to specifically cite the transparency achieved by use of PN. Figure-3.2 is a pictorial representation of the design synthesis to produce the product ‘A-... -M’. The steps leading to the evolution of a particular solution to the problem is visible to the human designer by means of such a pictorial diagram. Consider an example where it is required to build a table which can be easily moved, and the objective be the minimization of the total cost of the table. Let A, X, F, P, C and M in Figure-3.2 corresponds to table top, metallic frame, wooden frame, metallic legs, wooden legs and wheels respectively. The same is shown in Figure-3.3. By representing these chunks, pictorially in the PN gives the ability to the software to transparently and visually represent the steps which lead to a particular design solution, i.e., either ‘A-X-P-M’ or ‘A-C-F-M’. The rationale behind the choice of a particular solution is explained by the set of transitions which are fired, based on the rules specified (discussed in the following paragraphs).

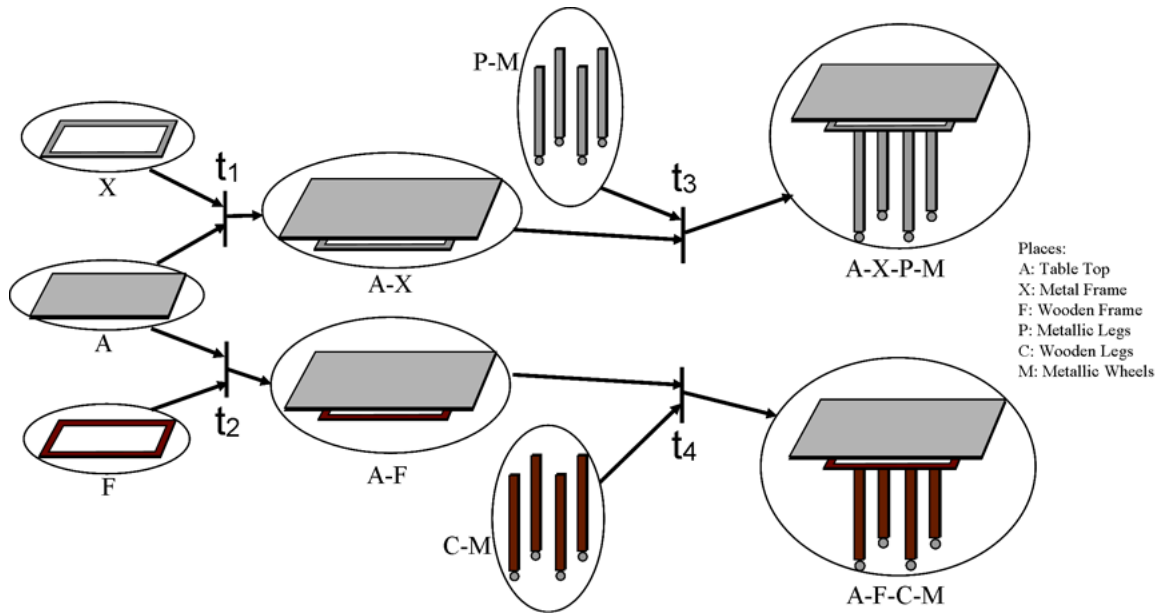


Figure 3.3: A sample table-design problem being solved and shown by use of the conceptualized Petri Net interface

At each stage of the design, the transitions that qualify for enabling are extracted and the choice of the transition to be fired is to be determined. This choice is not just dependent on the component cost but also its performance in minimizing the entire design cost in previously constructed designs. The comparison of transitions (firing a transition is equivalent to choosing the component at its source place) to be made is relative among the available options and is very vague in nature. A sample rule can be as follows: “Among all available transitions, if the cost of the component at the source place is least and its performance is highest, then the transition is fired”. Following this ideology, fuzzy logic principle (Zadeh 1975) is employed to decide between competing transitions. The fuzzy variables are identified as the transition’s cost and the transition’s preference. These fuzzy variables denote the cost (equal to the cost of the source chunk of the transition) and the preference value (tuple of the RB-PN) of the transition,

respectively. (The preference of the transition correlates to how well firing the transition (or choosing the component chunk at its source place) performed in previous designs; the preference of a transition is updated if it was fired to build a solution so that later instances can learn from experience. The mathematical update function of the preference value is listed consequently, in this section). The fuzzy sets are identified as being high, medium and low. To determine the association rules (trapezoidal) between the fuzzy variables and sets, a synthetic variable x is considered as shown in Table-3.2. The association functions of the synthetic variable x (which is equivalent to the association of the corresponding fuzzy variable) to the fuzzy sets have summarized in Table-3.3. It is to be noted here that the transition with low cost and high preference is desired (Table-3.2).

Fuzzy Variable	Synthetic variable (x)
The transition's cost	$x = \frac{(c_{max} - c)}{(c_{max} - c_{min})}$ <p>where,</p> <p>c = cost of the transition in consideration.</p> <p>c_{min} = the minimum cost among all the possible transitions.</p> <p>c_{max} = the maximum cost among all the possible transitions.</p>
The transition's preference	$x = \frac{(w - w_{min})}{(w_{max} - w_{min})}$ <p>where,</p> <p>w = preference of the transition in consideration</p> <p>w_{min} = the minimum preference among all the possible transitions</p> <p>w_{max} = the maximum preference among all the possible transitions</p>

Table 3.2: Conversion of the fuzzy variable to the synthetic variable x .

Range of a variable x	Association Function of the variable to the fuzzy sets:		
	High	Medium	Low
$0 \leq x < 0.25$	1.0	0.0	0.0
$0.25 \leq x < 0.4$	$(0.4 - x)/0.15$	$(x - 0.25)/0.15$	0.0
$0.4 \leq x < 0.6$	0.0	1.0	0.0
$0.6 \leq x < 0.75$	0.0	$(0.75 - x)/0.15$	$(x - 0.6) / 0.15$
$0.75 \leq x < 1.0$	0.0	0.0	1.0

Table 3.3: Association rules of x with the fuzzy sets high, medium and low.

The choice of the transition to fire is determined by a set of rules that constitute the design rationale of the software system. These rules are summarized in Table-3.4, and an example of a rule is constructed from a record of the rule-table is: “If the transition’s cost is low, and the transition’s preference is high, then $EF=1.00$ ”, there “EF” is the enabling factor of rule, which is used to obtain the modified truthfulness value of the rule:

$$\text{Modified truthfulness} = \text{Actual truthfulness} \times EF \quad (3-3)$$

For each transition, all of these rules are applied and the transition, which has the maximum modified truthfulness value, is fired by and the marking of the Petri net is adjusted.

Transition's cost	Transition's preference	Enabling Factor (EF)
Low	High	1.00
Low	Medium	0.90
Low	Low	0.15
Medium	High	0.80
Medium	Medium	0.75
Medium	Low	0.15
High	High	0.15
High	Medium	0.10
High	Low	0.05

Table 3.4: Rule-table of the software system, which represents the “design rationale”.

In the multi-agent system used in the proposed architecture, three agents are proposed – Configuration Agent (CA), Evaluator Agent (EA) and Fragment Agent (FA). The CA is responsible for the construction of the candidate solutions depending on the input and output specifications provided by the user. At each stage of the design, it extracts the transitions from the RB-PN that qualify for enabling, and depending upon the rules of the RB-PN, it fires transitions to advance the state of the design. At the end, it sends the completed candidate solution to the EA for evaluation. The EA evaluates the fitness of the solution by the objective function specified by the human designer and eventually, updates the weights of the transitions, which were fired to complete the solution, as per the following rule:

$$w = w_0 + \Delta w \quad (3-4)$$

$$\Delta w = \beta / C \quad (3-5)$$

where,

w_0 = initial preference value of the transition,

w = updated preference value of the transition,

β = update index ($0 < \beta \leq 1$) (specified by the human designer), and

C = Cost of the solution.

By updating the preference of the transitions, it is ensured that the agents learn from previous experience by use of the update preference in the fuzzy choice between the transitions. It is noted that the objective is to minimize the design cost of the entire solution, and hence the update rule is an inverse function of the solution cost; cheaper the solution, higher the preference. After updating the preference values in the RB-PN, the EA passes the solution to FA that optimizes the final solution considering each chunk as a fragment. For example, if 'A-B-C-D' is a solution ('A', 'B', 'C', and 'D' being chunks), and 'E' is another chunk matching the FPs of 'B' but having lower cost than 'B', then the solution 'A-E-C-D' is conceptualized (since the problem is to minimize the cost of the design) and the corresponding transitions' preference values are updated so that the CA can learn from this discovery during construction of the next candidate solution. The EA and the FA have a similar function, which is to update the transition preference values. However the FA has an additional responsibility of optimizing the solution by optimizing individual fragments of the solution. The interaction between the agents has been depicted in Figure-3.4. The model searches for the optimal solution

iteratively, given that the process of the construction, evaluation and chunk wise optimization of a candidate solution by the agents is considered as a single iteration.

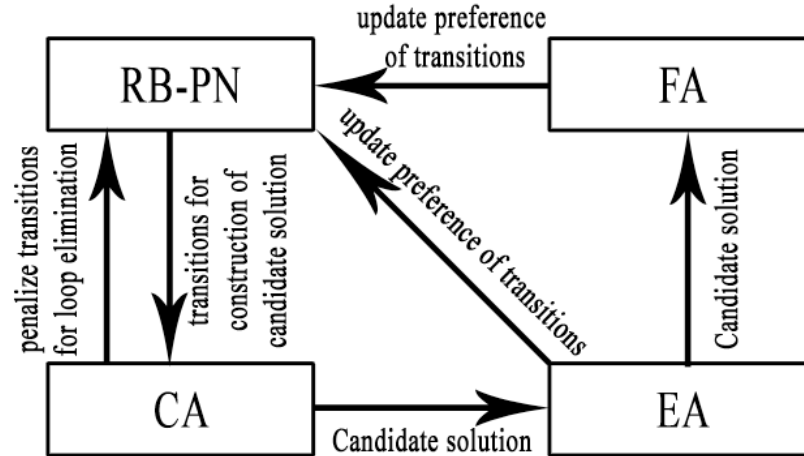


Figure 3.4: The Multi-Agent Structure.

It is to be noted that the while constructing a candidate solution, a part of the solution might be unnecessary loops of chunks, a worse situation being that the CA might get caught up in the loop and never complete the solution. To overcome this situation, the CA back-checks the solution state after the firing of each solution, to cut the loop and penalize the fired transitions (i.e., reduce the preferences of the transitions) which led to the creation of the loop, so they do not appear in later designs, as follows: If $t_1, t_2, t_3 \dots t_n$ were the transitions that were required to fire for the creation of a loop (in series), then the total penalty to be imposed is a fraction of the cumulative cost of the transitions. This penalty is distributed in parts to all the transitions ($t_1, t_2, t_3 \dots t_n$) in an arithmetic series, in which case t_1 and t_n are the least and most penalized transitions, respectively. The penalty ψ_i on the preference of the transition t_i is given by

$$\psi_i = \frac{i\Omega}{\left(\frac{n(n+1)}{2}\right)} \quad (3-6)$$

where

$$\Omega = \text{Total penalty} = \alpha \sum_{i=1}^n c_i$$

α = Penalty index ($0 < \alpha \leq 1$) (specified by the human designer)

i = Transition Index,

n = Total number of transitions in the loop, and

c_i = cost of transition t_i (equals cost of the source place of the transition).

Learning from design experience was demonstrated by Moss et al. (2004) by passing the learned chunks between different problems. This is similar to humans retaining good designs in memory and using them when they encounter a similar problem. In the current model, we propose the transfer of the transition preferences across problems for learning. It is to be noted that the preference of the transition are equivalent to the preferences of the source places, or the chunks, of the transition. Hence, passing the preferences effectively infers the passing of the knowledge of the good and bad chunks. The pseudo code of the architecture is presented in Table-3.5 and Figure-3.5, which illustrates the flow chart of the model.

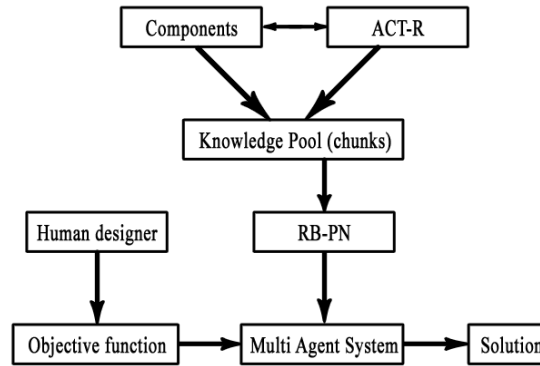


Figure 3.5: Flow chart of the architecture.

```

Human designer initializes the parameters
For each iteration i
Begin
  While solution S is incomplete
  Begin
    CA fires transitions based of the RB-PN to complete S
    If any loops have been encountered
    Begin
      CA eliminates the loop
      CA penalizes the transitions which lead the creation of the loop
    End
  End
  End
  CA sends S to Evaluator Agent (EA)
  EA updates the transition weights related to S and passes S to Fragment Agent (FA)
  FA optimizes S chunk-wise to generate solution So
  FA updates the transition weights corresponding to So
End
Print the best solution obtained
  
```

Table 3.5: Pseudo Code.

3.4 Example, results and discussion

The requirements for a human-centered CAD system proposed in the current work are revisited as having transparency, design rationale and cognitive capabilities. The proposed architecture described previously is hypothesized to be able to lead to a software system that can fulfill these requirements.

Petri nets and fuzzy logic together satisfy the transparency and design rationale requirements as fuzzy logic gives a natural way of decision making which can be easily understood by human designers and by means of the pictorial representation of the design synthesis process (an example of which is shown in Figure-3.3). The pictorial representation of the transitions fired to complete the design provides a transparent and white-box view of the actions executed by the agent system at every step of the design synthesis.

The capability of cognition is hypothesized to be achieved by use of the ACT-R architecture of Moss et al. (2004). This hypothesis is verified in this section as the original architecture of (Moss et al. 2004) is not used 'as is' but is modified and a rule based Petri Net is used as the procedural memory of the ACT-R. Some other issues related to the architecture such as convergence and loop elimination by the multi-agent system are also tested in the implementation.

To determine the effectiveness of the proposed agent system for supporting engineering design, a design example is essential as long as a sufficiently large component base is available for the design synthesis. For example, in the case of solving electro-mechanical design problems, the component base, which primarily consists of

elements such as footpad, lever, springs, pulleys, belts, cylinders, generators, resistors, etc., has to be large enough so that the software agents be able to not only produce variants of design solutions using different combinations and configurations of components, but also be able to learn which components yield better design solutions and which effect the design solution the worst, with respect to the objective function (say minimization of the design cost). Table-3.6 depicts a chunk (belt-pulley chunk) as per the proposed architecture.

Any electro-mechanical problem can be broken down by the human designer as a set of input and output functions. For example, a pressure-gauge problem has a pressure source as an input and the output is a dial display, which displays the pressure applied by the source. Hence, the human designer instructs with the agent to produce a design solution which has the input functional parameters similar to that of a cylinder and the output functional parameters similar to that of a dial, using the electro-mechanical component base. These functional parameters can be labeled as $(D1,I1)$ and $(D2,I2)$, where $D1,D2$ represent the domains of the input (cylinder) and output (dial) and $I1,I2$ represent the interfaces of the input (cylinder) and output (dial) respectively. It is noted here that the various domains and interfaces of electromechanical components can be labeled, in the software system, as two separate lists: $(D1,D2,D3, \dots)$ and $(I1,I2,I3, \dots)$, and each component's functional parameters can be labeled as the combination of two labels from the lists (as shown above).

Labeling the various domains and interfaces offers the advantage of testing the software system even in the case of lack of sufficient data of machine design components, as hypothetical combinations of domains and interfaces can be created and

tested. Also, such an arrangement does not hamper the design synthesis of the software system. Table-3.7 presents the same example of the design chunk as shown in Table-3.6, but using hypothetical domains and interfaces. In the present article, due to the lack of sufficient data of machine design components, a computational domain is constructed as follows: seven hypothetical domains are taken into consideration (D1, D2, D3...D7) which represent domains like translation, rotation etc. Similarly for the interfaces, which in general represent components like bolt, shaft-hole etc, seven hypothetical interfaces (I1, I2, I3...I7) are considered. Using this notation, a sample chunk has already been depicted in Table-3.7.

Chunk Belt-pulley	
Isa	: Design chunk
Input Domain	: Translation
Input Interface	: Bolt
Output Domain	: Rotation
Output Interface	: Shaft-hole
Components	: Belt, pulley
Connectivity	: Port-1 of Component-1 is connected to Port-1 of Component-2
Goodness	: 180

Table 3.6: An example design chunk (adapted from Moss et al. (2004)).

As shown in Table-3.7, chunk #16 is an ACT-R design chunk. It has the domains D4, D1 for input and output, and interfaces I5, I6 for input and output, respectively. This particular chunk has two components C1, C2 which are obtained from a predefined catalogue. The cost of the chunk is determined by the sum of the cost of the components

constituting it. The model was tested for a pool of 5000 chunks, which collectively represent the knowledge pool. A chunk string, ‘#12-#54-#93’, represents a design solution with the specifications: input FPs of chunk ‘#12’ and output FPs of chunk ‘#93’.

Design-Chunk #16	
Isa	: Design-chunk
Input-domain	: D4
Input-Interface	: I7
Output-domain	: D8
Output-Interface	: I3
Components	: C1, C2, etc.
Connectivity	: Port-1 of C1 is connected to Port-1 of C2.
Goodness	: 180

Table 3.7: A sample design chunk as represented in the model.

#4037-#85
#4037-#85-#614
#4037-#85-#614-#524
#4037-#85-#614-#524-#85
#4037-#85
#4037-#85-#679

Table 3.8: An example of the elimination of loops by the CA.

The capability of the CA to overcome loops was tested. It is desired that loops are not only removed as and when they are removed, but also the system does not repeat the mistake again – a learning mechanism comes into play here. A sample solution

generation process has been shown in Table-3.8 which illustrates that loop ‘#85-#614-#524-#85’ was encountered and was later not again repeated by penalizing the transitions which led to the formation of the loop. This component-loop formed because the chunk ‘#85’ was able to connect to the chunks ‘#614’ and ‘#524’ through its output and input functional parameters respectively. Several of such loops (of different configurations) were encountered and the system was able to not repeat the formation of each of the previously encountered loops and hence was able to increase the efficiency of producing design solutions using memory.

The model was tested for convergence by solving a hypothetical problem to produce a design with having a specified combination of the input and output FPs. For each design problem, the model was run for 200 iterations and the results (cost of the solutions) of a randomly selected design problem have been plotted in Figure-3.6. The cost converges after around 92 iterations.

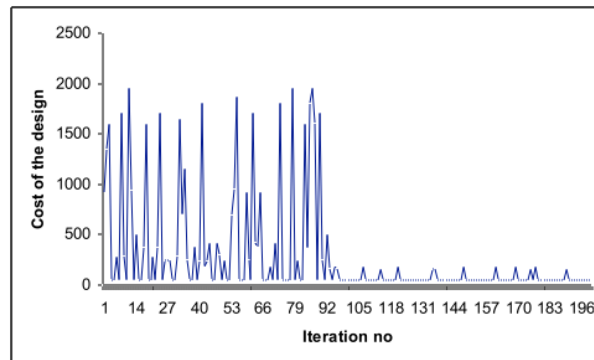


Figure 3.6: Cost of the transitions obtained in the iterations.

The model was also tested for the knowledge transfer ability by passing the transition information (the weights) and the model was once again run for 200 iterations. The cost of the solutions obtained in the iterations is plotted in Figure-3.7. The average

cost of the design is calculated for both run (with and without transition knowledge) and the results are displayed in Figure-3.8. As is clearly visible from all these graphs, the model converges in about 60 iterations when knowledge transfer is used – 30% less number of iterations.

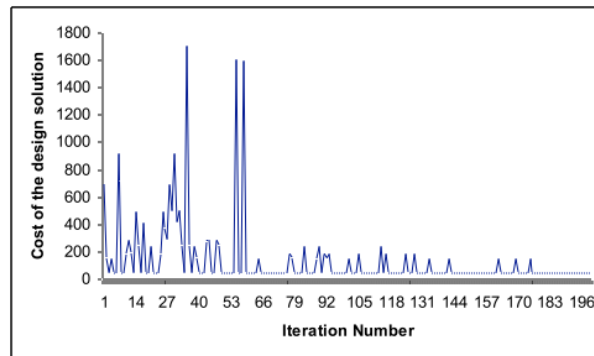


Figure 3.7: Cost of the designs obtained with prior transition knowledge.

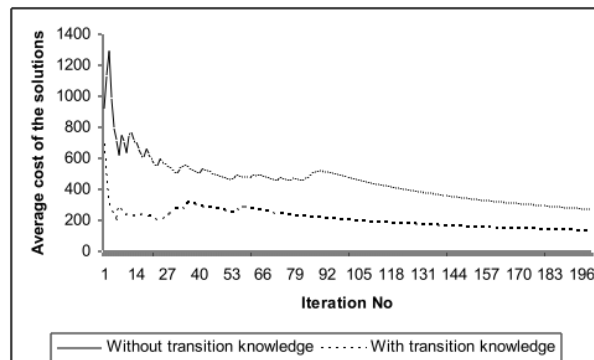


Figure 3.8: Average cost of the solutions obtained with and without transition knowledge.

3.5 Limitations of the proposed architecture

Though the model is proposed for the minimization of an objective function, it can be used to maximize an objective by adjusting the conversion of the fuzzy variable (the transition) to the synthetic variable (x) and the weight update function of the agents.

A limitation of the proposed architecture is that the objective is to optimize the design

from the single perspective (in this case, the cost of the design). The architecture can be extended to account for multiple objectives by using approaches such as Pareto-optimality (Deb 2001), which can extend the single objective approach to the multi objective approach. The proposed architecture is described for the context of problems, which can be characterized as optimization of the design. The architecture can be extended to other contexts by doing appropriate modifications in the representation of the chunks and the rule based PN.

The tools used in the proposed architecture help to impart the desired capabilities very effectively. However, they induce a few limitations as well. The use of Petri nets to represent the design synthesis leads to complex pictorial representations because of the immense number of places and transitions in the net (Murata 1989). Since this is a classical problem with Petri nets, several studies in literature can be found which deal with this problem by using approaches such as hierarchical PNs (Huber et al. 1990).

3.5 Conclusion

There are several salient advantages behind this architecture. First, Petri net best describes a complex event-driven dynamic process, which is the nature of engineering design. Use of the Petri net helps to impart the capability of transparency to the system. Second, ACT-R is a very good formalism for constructing the agent to simulate the human cognitive process of problem solving. Third, fuzzy logic is able to represent knowledge and logic in a natural way and is further convertible to artificial neural networks (Pedrycz and Gomide 1999). Finally, the proposed architecture is built upon a multi-agent system which builds a strong foundation for future development of an effective interface between the human designer and the computer designer. Experiments

have shown the model to yield satisfactory results in eliminating the encountered loops from the design synthesis, the convergence of the model and the transfer of knowledge.

It can be envisioned that several architectures can be built around the principles of the capabilities mentioned in this paper. The differences would be in the tools used and the representation of information in these architectures. For example, one could use the QN-MHP for cognitive modeling (Wu et al. 2008). A human designer making decisions of the choice of the components can replace the fuzzy production rules. However, it is clear that all these architectures would depend on the context of the CAD system, the cooperation between the human designer and the CAD agents, etc. At their core, it is evident that the capabilities discussed provide a strong foundation for HAS for CAD.

In essence, a basic understanding of the HAS is extremely vital. This case study about HAS for a CAD agent system is the first study to be conducted from the viewpoint of core capabilities discussed in this thesis. Several other case studies have to be studied to complete the list of the capabilities desirable for better HMI.

Chapter 4: An integrated framework for human state inference

4.1 Introduction

This chapter focuses on development and testing of a new framework ideology for human state inference. As discussed in Chapter-2 of this thesis, an integrated framework is necessary to be developed because none of the experts is powerful enough to infer the state for every situation (Gray 2007).

The framework is socially inspired. Consider an example where a problem is given to a group of people in a community. The people divide themselves into various sub-groups and look at the problem from various perspectives. Each person has a slightly different way (methodology) to solve the problem, which is based on his/her educational background. People also differ in their efficiency (experience) to solve a problem. For example, a mathematician is more likely to solve a numerical problem with more efficiency when compared to an art major. Based on his/her methodology and experience, each person gives his/her own opinion of the candidate solution of the problem. If the problem is extremely challenging, integrating the opinions from various people to collectively solve the problem from various perspectives could yield a good result, due to the diverse methodologies, experiences and perspectives. In this thesis, it is attempted to mimic the above-discussed social behavior to solve the problem of inferring the human state. Several algorithms (which represent the experts) and signals from sensors (which denote the perspectives) are modeled in the framework.

4.2 Proposed framework

For the purpose of explaining the framework, the following notations are used: Let C_i denote the i^{th} cue, which can be independently used by an expert to infer the human state and m denote the total number of such cues. Let E_j be the j^{th} category of expert and n denote the total number of the categories of experts. Let O_{ij} be the estimate (or opinion) of the mental state inferred by the j^{th} expert using the i^{th} cue.

The experts represent the technique of inferring the mental state and each expert's opinion is a potential estimate of the state. The methodology of integrating the opinions of the experts to obtain the combined estimate of the mental state represents the procedure of integration. There can be several procedures for integrating the opinions with various semantics and hierarchies. However, this thesis is restricted to four procedures originating from different schools of thought of community behavior of problem solving. Let OP_k denote the opinion of the k^{th} procedure of integration.

4.2.1 Procedure I

In this procedure, the opinions of the experts of the same category are combined in the first place to represent the shared opinion of the category of experts. This gives the chance to the experts with a similar problem solving technique (category) to view the problem from all perspectives (cues). The opinions from all the categories of experts are then combined to obtain the estimate representative of the entire community at large.

As shown in Figure-4.1, this procedure consists of two levels. In the first level, the opinions from all the experts of the same category are fused together to obtain the combined opinion OE_j , which represents the opinion from the j^{th} category of experts. In

the second level, the opinions OE_j from the first level are fused together to obtain the final estimate of the mental state, which is denoted by OP_1 .

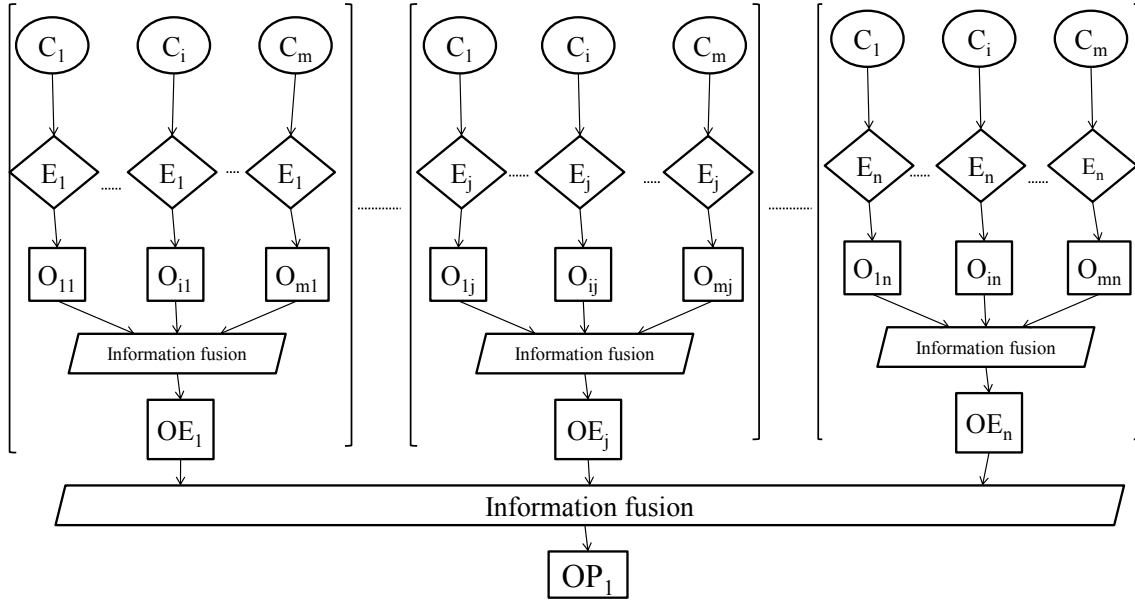


Figure 4.1: Framework of Procedure I. The problem is first viewed by the same category of experts from all perspectives.

4.2.2 Procedure II

In this procedure, the opinions from the experts, which infer from the same cue, are combined in the first place to represent the estimate from the particular perspective (cue). This gives the chance to obtain the estimate from the same perspective of the problem (cue) but combining the various techniques (experts). The opinions from all the perspectives (cues) are then combined to obtain the estimate representative of the entire community at large.

As shown in Figure-4.2, this procedure consists of two levels. In the first level, the opinions from all the experts using the same cue are fused together to obtain the

combined opinion OC_i that represents the opinion from the i^{th} cue. In the second level, the opinions OC_i from the first level are fused together to obtain the final estimate of the mental state, which is denoted by OP_2 .

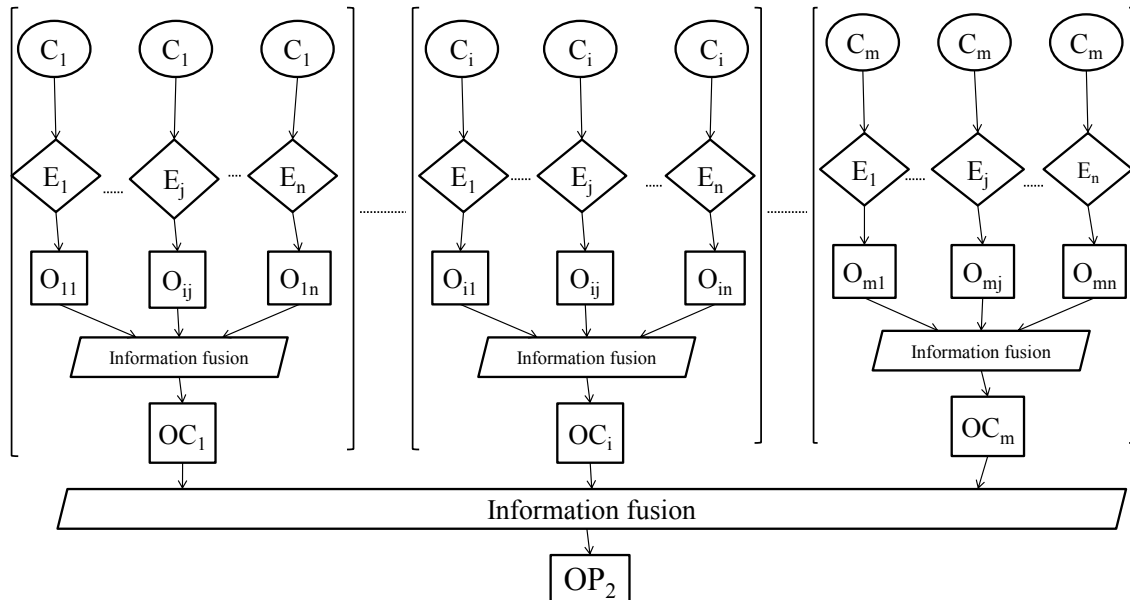


Figure 4.2: Framework of Procedure II. The problem is first viewed by all experts from the same perspective (cue).

4.2.3 Procedure III

In this procedure, all opinions are combined with a minimal hierarchical structure to result in an opinion representing the entire community.

As shown in Figure-4.3, this procedure consists of only one level. The opinions from all the experts O_{ij} are fused together to obtain the final estimate of the mental state, which is denoted by OP_3 .

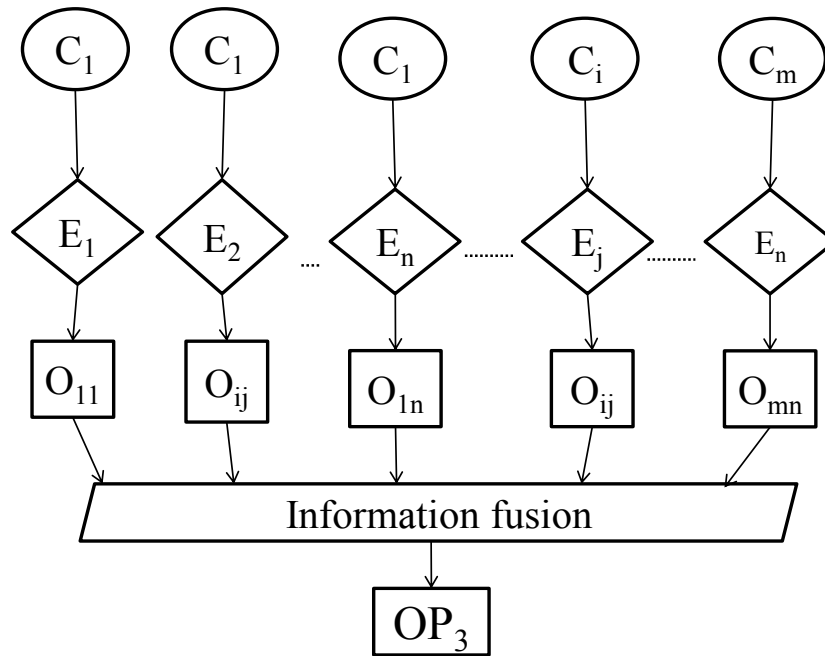


Figure 4.3: Framework of Procedure III. The opinions are fused together without any hierarchy

4.2.4 Procedure IV

In this procedure, the opinions obtained from other procedures are fused together in a hope to capture the benefits of various structures (or semantics) of the procedures.

As shown in Figure-4.4, at the conceptual view, this procedure consists of only one level. The opinions from all the other procedures OP_k ($k \neq 4$) fused together to obtain the final estimate of the mental state.

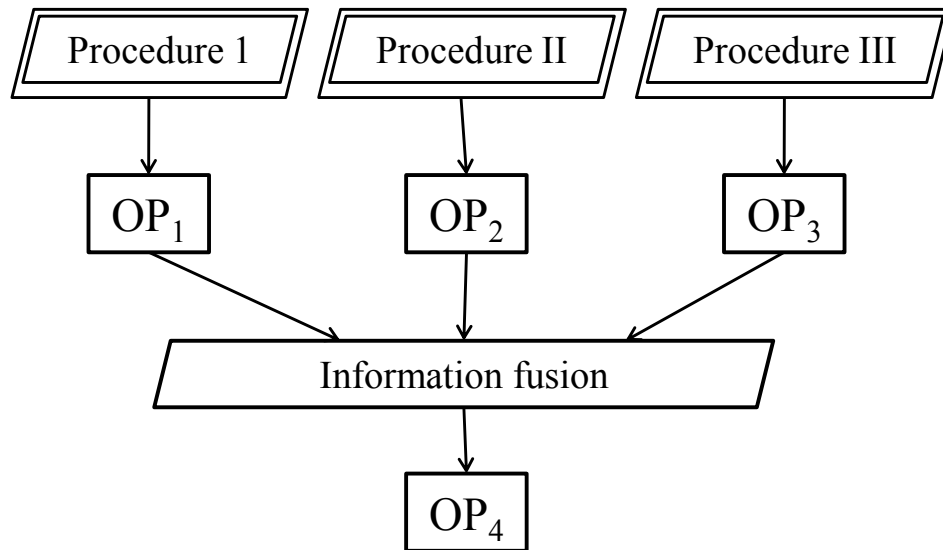


Figure 4.4: Framework of Procedure IV. The opinions from other procedures are fused together.

4.2.5 Information fusion

In the aforementioned procedures, the opinions received from various experts can be fused together by any information fusion technique such as Ordered Weights Aggregation (OWA) (Yager 2004), expert opinion elicitation (Clemen and Winkler 1999). In general, these techniques differ in terms of the underlying weighing function.

4.3 Implementation

4.3.1 Example

Yang et al. (2008) presented a fatigue inference system based on the principle of multi-modality and information fusion. The cues used to infer the fatigue of the drivers are as follows:

- 1) Sleep Quality (SQ)

- 2) Driving hours (DH)
- 3) Electroencephalogram (EEG)
- 4) Electrocardiograph (ECG)
- 5) Eye movement (EM)

Quality of the sleep (SQ) and number of driving hours (DH) are considered to be the causal features of fatigue in human drivers. EEG and ECG are means to measure the electrical activity of the brain and the heart respectively, which are argued to contain information about fatigue. Percent eye closure (PERCLOS) was taken as the eye movement activity (EM) which contains the information about fatigue.

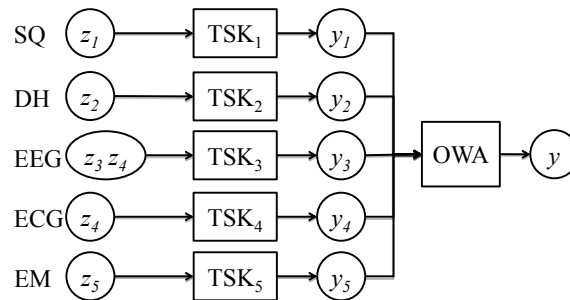


Figure 4.5: Fatigue inference model proposed by Yang et al. (2008)

In their work, raw data collected from various sensors from 9 human drivers were normalized and pre-processed to yield six variables z_i ($i = \{1... 6\}$) which are representative of the five cues. These variable had values on a continuous numerical scale between 0 and 1. Five Takagi-Sugeno-Kang (TSK) based experts, one for each cue as shown in Figure-4.5, were developed to infer the fatigue value, y_j ($j = \{1... 5\}$), representative of each cue, which were also on a continuous numerical scale between 0 and 1. These fatigue estimates were then fused together using the principle of ordered

weighted aggregation (OWA) (Yager 2004) to get the final value of the fatigue state (γ) of the human drivers.

In this thesis, fatigue is considered to be the human state of particular interest. The architecture of Yang et al. (2008) can be generalized to suit any other type of expert by replacing the TSK technique with other technique. For the purpose of implementation of the proposed framework, the same dataset as Yang et al. (2008) is used (Appendix-B).

4.3.2 Experts used

In order to test the framework, it is necessary to collect experts in inferring fatigue state. Various inference techniques can be employed on the dataset to generate different types of experts. In this thesis, four types of experts are used to infer the fatigue, and they are introduced as follows:

- 1) Artificial Neural Network (ANN) based experts
- 2) Statistical (STAT) experts
- 3) Takagi-Sugeno-Kang (TSK) based experts
- 4) Static Bayesian Network (SBN) based experts

Each of the above experts uses the multi-modality structure with OWA (see Appendix-C) as the information fusion technique as discussed in (Yang et al. 2008). ANN based experts are trained using the data set. STAT experts were constructed using regression in SPSS. TSK based experts were used as provided by the authors of Yang et al. (2008). Based on the experience of the authors, prior probabilities were assigned to the SBN based experts.

4.3.2.1 Statistical experts (STAT)

The following statistical models were obtained after conducting linear regression analysis on the data presented in Appendix-B.

The statistical model for inferring fatigue (y_1) from the sleeping quality (z_1) is as follows:

$$y_1 = 1.0034 - 1.0006z_1 \quad (4-1)$$

The statistical model for estimating fatigue (y_2) from driving hours (z_2) is as follows:

$$y_2 = 0.0429 + 0.9408z_2 \quad (4-2)$$

The statistical model for the fatigue value (y_3) from the EEG analysis (z_3 and z_4) was obtained from a defined matrix as follows.

$$y_3 = \begin{bmatrix} 0.01 & 0.05 & 0.10 & 0.20 & 0.30 \\ 0.05 & 0.45 & 0.49 & 0.50 & 0.55 \\ 0.10 & 0.49 & 0.55 & 0.60 & 0.65 \\ 0.20 & 0.50 & 0.60 & 0.75 & 0.80 \\ 0.30 & 0.55 & 0.65 & 0.80 & 0.91 \end{bmatrix} \quad (4-3)$$

The row (m) and column (n) number corresponding to signals z_3 and z_4 respectively are obtained as follows.

$$m(n) = \left\langle \begin{array}{l|l} 1 & 0 \leq z_1(z_2) < 0.4 \\ 2 & 0.4 \leq z_1(z_2) < 0.5 \\ 3 & 0.5 \leq z_1(z_2) < 0.6 \\ 4 & 0.6 \leq z_1(z_2) < 0.75 \\ 5 & 0.75 \leq z_1(z_2) \leq 1 \end{array} \right\rangle \quad (4-4)$$

The statistical model for the fatigue value (y_4) from the ECG analysis (z_5) was obtained from a defined matrix as follows.

$$y_4 = \left\langle \begin{array}{l|l} 0.5889z_5 + 0.0337 & 0 \leq z_5 < 0.41 \\ 5.75z_5 - 2.1075 & 0.41 \leq z_5 < 0.45 \\ 0.2843z_5 + 0.3586 & 0.45 \leq z_5 < 0.667 \\ 4.3373z_5 - 2.343 & 0.667 \leq z_5 < 0.75 \\ 0.0763z_5 + 0.8486 & 0.75 \leq z_5 \leq 1 \end{array} \right\rangle \quad (4-5)$$

The statistical model for the fatigue value (y_5) from the eye movement analysis (z_6) was obtained from a defined matrix as follows.

$$y_5 = \left\langle \begin{array}{l|l} 1.2048z_6 & 0 \leq z_6 < 0.083 \\ 0.0753\exp(1.689z_6) & 0.083 \leq z_6 < 0.667 \\ 7.3494z_6 - 4.642 & 0.667 \leq z_6 < 0.75 \\ 0.2767z_6 + 0.6533 & 0.75 \leq z_6 \leq 1 \end{array} \right\rangle \quad (4-6)$$

The aforementioned statistical models for estimating the fatigue levels are obtained by analyzing the data of Yang et al. (2008). The variables z_i ($i = \{1... 6\}$) and y_j ($j = \{1... 5\}$) are representative of the various cues and the fatigue values.

4.3.2.2 Artificial Neural Network based experts (ANN)

The data from Yang et al. (2008) was used as training data for training five neural networks (one for each mode of cues). The MATLAB source codes for the networks were downloaded from collection of open-source codes from Phil Brierley's website (<http://philbrierley.com>) which were modified to suit the current implementation. The networks consisted of one input (two inputs in case of EEG analysis) and one output (fatigue). One layer of 3 hidden neurons was implemented.

4.3.3.3 Static Bayesian Network based experts (SBN)

The Static Bayesian Network based experts for inferring fatigue from the various cues were constructed by using an approximate guess of the prior probabilities. There are all simple cause-effect links with individual cues as the cause and fatigue state as the

effect. The guesses of the prior probabilities were made by manually modifying the prior probability values until an acceptable accuracy of inference was achieved (see Table-4.1). As these values were not taken from any previous studies in literature, it is not surprising that the SBN based experts performed the poorest in inference.

4.3.3.4 Takagi-Sugeno-Kang based experts (TSK)

The TSK based experts were used ‘as reported’ by Yang et al. (2008). The source codes were obtained from the authors of Yang et al. (2008).

4.3.3.5 Performance of individual experts

The experts were individually used to infer the fatigue. Same data set was presented to each of the experts and the inferred values were recorded. This process was repeated five times and the final summary of the results is shown in Table-4.1. The mean of error (eq. 4-8) and variance in error (eq. 4-9) of inferring the fatigue state for the given dataset is reported in Table-4.1, which is representative of all the five runs.

$$e_i = \left(\frac{y_i^{predicted} - y_i^{actual}}{y_i^{actual}} \right) \times 100 \quad (4-7)$$

$$Mean\ of\ error = e_{mean} = \frac{\sum_{i=1}^N e_i}{N} \quad (4-8)$$

$$Variance\ in\ error = e_{var} = \frac{\sum_{i=1}^N (e_i - e_{mean})^2}{N} \quad (4-9)$$

The experts are ranked in Table-4.1 by their accuracy in inferring the fatigue levels. The ANN experts performed the best with the lowest mean error and lowest variance in the error of inference. The performance of the STAT experts is dependent on the models used and the estimated models were found to yield satisfactory results. The

TSK experts were used ‘as is’ from Yang et al. (2008) and they resulted in an 8.19% mean of error and 265.50% variance in error (about 10% mean error was reported in the original paper). The SBN experts performed the poorest amongst the four experts because of the prior probability values given to them.

Rank of the expert	Type of expert	Mean of error (%)	Variance in error (%)
1	ANN	3.06	90.14
2	STAT	4.68	164.06
3	TSK	8.19	265.50
4	SBN	10.28	446.15

Table 4.1: Mean and variance of the error in inference by individual expert for five runs

It is evident that the experts considered are quite diverse in terms of both the principle and the performance of inferring the fatigue. This closely mimics the diversity in a community: each person has his/her method and efficiency in solving a problem.

4.3.3 Performance of the framework

With four types of experts, several subgroups of experts can be formed and their inference results can be compared to an individual expert and other sub-groups. In this thesis, the scope is limited to comparing the performance of the various subgroups of experts versus the TSK experts of Yang et al. (2008). This narrows the combinations to a lower count as reported in Table-4.2 in which each combination has either the TSK experts or the experts ranked below namely the SBN experts. The excluded combinations are of the experts ranked higher than the TSK experts and comparison makes little sense in such cases.

Each of the combination of experts in Table-4.2 was used to infer the fatigue state using the framework discussed in the aforementioned section. Their performances are tabulated in Table-4.3 in which the mean and variances in the errors of inference by the combination of experts and the procedure are presented. The results are the outcome of repeating the experiment three times.

Name	Number of experts	Types of experts
C-4-1	4	ANN, STAT, TSK, SBN
C-3-1	3	ANN, STAT, TSK
C-3-2	3	ANN, STAT, SBN
C-3-3	3	STAT, TSK, SBN
C-2-1	2	ANN, TSK
C-2-2	2	ANN, SBN
C-2-3	2	STAT, TSK
C-2-4	2	STAT, SBN
C-2-5	2	TSK, SBN

Table 4.2: Combination of experts

Name	Procedure I	Procedure II	Procedure III	Procedure IV
	Mean (Var)	Mean (Var)	Mean (Var)	Mean (Var)
C-4-1	6.20 (161.32)	7.04 (208.35)	6.56 (167.23)	6.64 (162.43)
C-3-1	4.70 (113.12)	2.88 (34.22)	4.54 (100.46)	4.01 (73.34)
C-3-2	6.91 (203.32)	3.59 (94.74)	7.69 (226.30)	6.27 (168.17)
C-3-3	6.72 (175.63)	4.13 (79.66)	6.94 (183.52)	5.88 (133.05)
C-2-1	6.41 (167.25)	1.03 (32.16)	7.49 (193.49)	4.91 (103.58)
C-2-2	6.91 (209.12)	2.91 (89.99)	7.77 (244.68)	5.97 (164.32)
C-2-3	6.02 (147.8)	0.68 (18.13)	5.54 (133.54)	3.88 (68.52)
C-2-4	8.06 (256.29)	2.72 (60.98)	7.37 (210.63)	6.09 (149.79)
C-2-5	10.05 (336.63)	2.63 (76.24)	8.81 (310.84)	6.97 (198.18)

Table 4.3: Mean and variance of the error (%) in inference by the various combinations for three runs

4.3.4 Observations

The following observations can be made from Table-4.2 and Table-4.3:

First, the best result (lowest mean error and variance in error) in all combinations was achieved in Procedure II with the exception of the combination C-4-1. This finding makes us believe that when a group of two to three categories of experts collaborate to solve a problem, best results are achieved when the problem is looked from one perspective at a time with all experts giving their opinion. Later, the opinions from all perspectives should be integrated. It is also worthwhile to note that the choice of the experts plays a significant role in the inference accuracy of the combination and hence

the above finding may only be applicable to the tested set of experts and procedures using the dataset.

Second, it is worthwhile to note that given a piece of information (i.e., data set), the result is dependent on not only the technique used (i.e., the expert) but also the procedure of using these techniques (i.e., the method of integration). In all the combinations, the results are significantly different in each of the procedures. A paired sample student's t-test shows that there is a significant difference in the means of the errors for most of the procedures (95% confidence interval). However no significant difference could be found between Procedure 1 and Procedure 3 ($p=0.764$).

Third, the combination C-2-5, which is composed of TSK and SBN experts, gives a better result in Procedure II and Procedure IV when compared to the results obtained by just the TSK experts, even though SBN is a lower ranked expert than TSK. This performance improvement can be explained by the effect of diversity in a community. Experts or people have pros and cons in their problem solving techniques and working in a collaborative manner gives the advantages of using the pros of the techniques.

In general, communities of high ranked experts yield a better result when compared to the other combinations. For example, the combinations C-3-1 (Ranks: 1, 2, 3) yield a lower mean and variance in error when compared to C-3-2 (Ranks: 1, 2, 4) which in turn gives better results when compared to C-3-3 (Ranks: 2, 3, 4). This can be explained by the high collective intellectual capacity of a group when the individual members are good experts.

Interestingly, the best result is obtained by the combination C-2-3 (i.e., STAT and TSK based experts). In particular, Procedure II for this combination yields a mean error

of 0.68% and a variance in error of 18.13%. At this point, this achievement could not be explained.

4.4 Conclusion

In this chapter, a framework to integrate various algorithms to infer the human state using various cues was presented. The framework was implemented using four types of algorithms to infer the fatigue state of human drivers.

For the proposed set of experts, cues, procedures and dataset, it is concluded that using several algorithms to infer the human state is significantly better than using one algorithm only owing to the diversity in the technique of inference. In most cases, the best results were obtained when a problem is investigated by each perspective by all experts first and then fusing the results obtained from all perspectives (i.e., Procedure II). It has also been found that a combination of high ranked experts yields a better result than other combinations owing to good experiences of the experts in inference. It is also worthwhile to note the conclusion in Yang et al. (2008) that several cues yield higher accuracy in inference of the human state.

In this study, OWA (Yager 2004) was used as an information fusion technique for the current implementation. A study of the influence of different information fusion techniques could not be done in the present work. A future study seems to be necessary to study the influence of different information fusion techniques on the inference.

One of the biggest challenges that were faced was the “ground-truth” of the fatigue state value. As with almost all cases of human state detection, there is no “real value” of the state. However to train any algorithm to infer the fatigue state or other

human states such as workload or emotion and so on, it is vital to have a good set of input and output data. Different research studies report various ways for establishing a ground-truth value for their evaluation purposes. The approach is influenced by not only the researcher's data processing technique but also the experimental context. For this reason, it is extremely challenging to compare two studies unless they have the same dataset, data processing/signal smoothing technique and experimental context. The results obtained from this framework could only be compared to results from a previously reported study of TSK experts (Yang et al. 2008) for the aforementioned reason. However, it is extremely necessary and urgent for the research community to formulate an acceptable standard(s) for the "ground-truth" value of human state.

Future studies of implementation of the framework to other human state such as emotion, mental workload etc., are essential to further advance this direction of research. In light of the variation in the accuracy of inference by changing the experts and procedures, the proposed framework for human state inference appears to be quite promising.

Chapter 5: Effect of representation of machine's state on user's actions

5.1 Introduction

This chapter presents the work with respect to the third objective of this thesis study. In Chapter-2, background was discussed about the paradigm of representation of machine's state in a natural way. There can be potential benefits of such a representation in HMI. The hypothesis is that there is an effect on the user's actions in a human-machine systems task, when the machine's state is presented in a natural manner.

In order to test the above hypothesis, experiments were carried out on a simulated test-bed to collect data for statistical analysis. A detailed description of the test-bed and the experiment design is provided in Section-5.2. The results from the statistical analysis of the data and their implications are given in Section-5.3. The chapter is concluded in Section-5.4.

5.2 Simulated test-bed

To design and develop a test-bed to carry experiments and collect data, a car was chosen as the machine to be considered. Human-vehicle interaction is a classic case of human-machine interaction. In the scope of this thesis study, only one machine state variable is implemented, which is the available fuel in the vehicle. This state was chosen because the change in fuel level is a frequent occurrence. The lack of fuel in the vehicle is associated to it being 'thirsty' (Norman 1992). 'Thirst' is a state of the human beings that is always taken care of by drinking water. It seems logical that as communication

between humans and vehicles evolve, the natural way that the car communicated to the human being about its need for fuel would be that it is thirsty. However, other designers might not use this linkage between fuel level and ‘thirst’ label. Fuel is a source of energy to car as food is to human beings. Following this notion, lack of fuel may be labeled as the car being ‘hungry’. However in this scenario, it was felt that ‘thirst’ might be more appropriate label by virtue of the liquid state of the fuel. The choice of the label is a design decision and is subject to the opinions of the designers.

Following this idea, a simulated test-bed is created to represent the fuel state in the natural way to the human being. The hardware of the test-bed is a Macbook Pro. The software behind the test-bed is five versions of a game interface, which is custom-made using the JAVA programming language. The game is a driving game in which the user drives a vehicle on a path using the controls on the keyboard.

5.2.1 Basic components of the interface

There are seven components of the interface display of the game as shown in Figure-5.1.

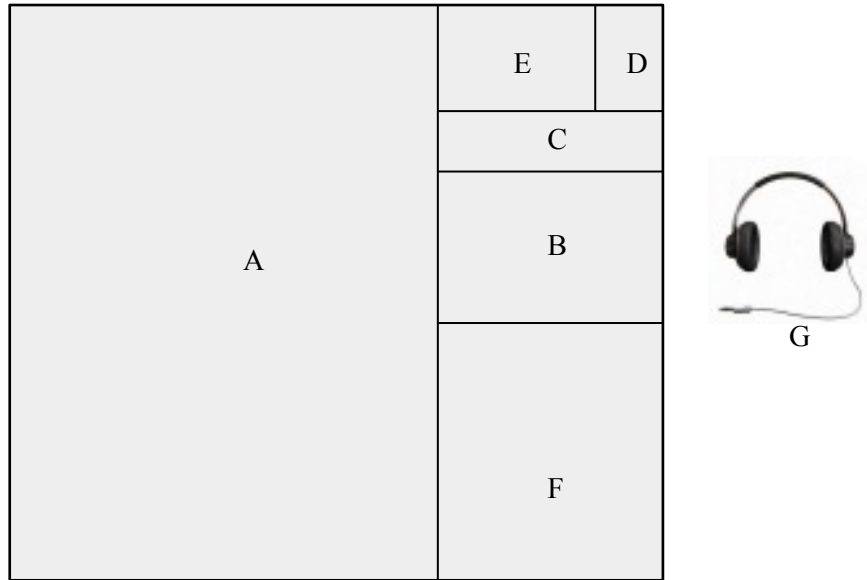


Figure 5.1: Components of the interface of the test-bed

5.2.1.1 Component A

This area/component displays the map on which the user drives the car by use of arrow keys on the keyboard. As shown in Figure-5.2, the vehicle is represented by a red box on the screen, which is driven on a specified path, marked by black lines. As the user navigates the vehicle on the path, the vehicle loses fuel. The user can refuel the car at any fuel station on the map as indicated. The location of the fuel stations along the path were determined by repeated testing of the prototype interface.

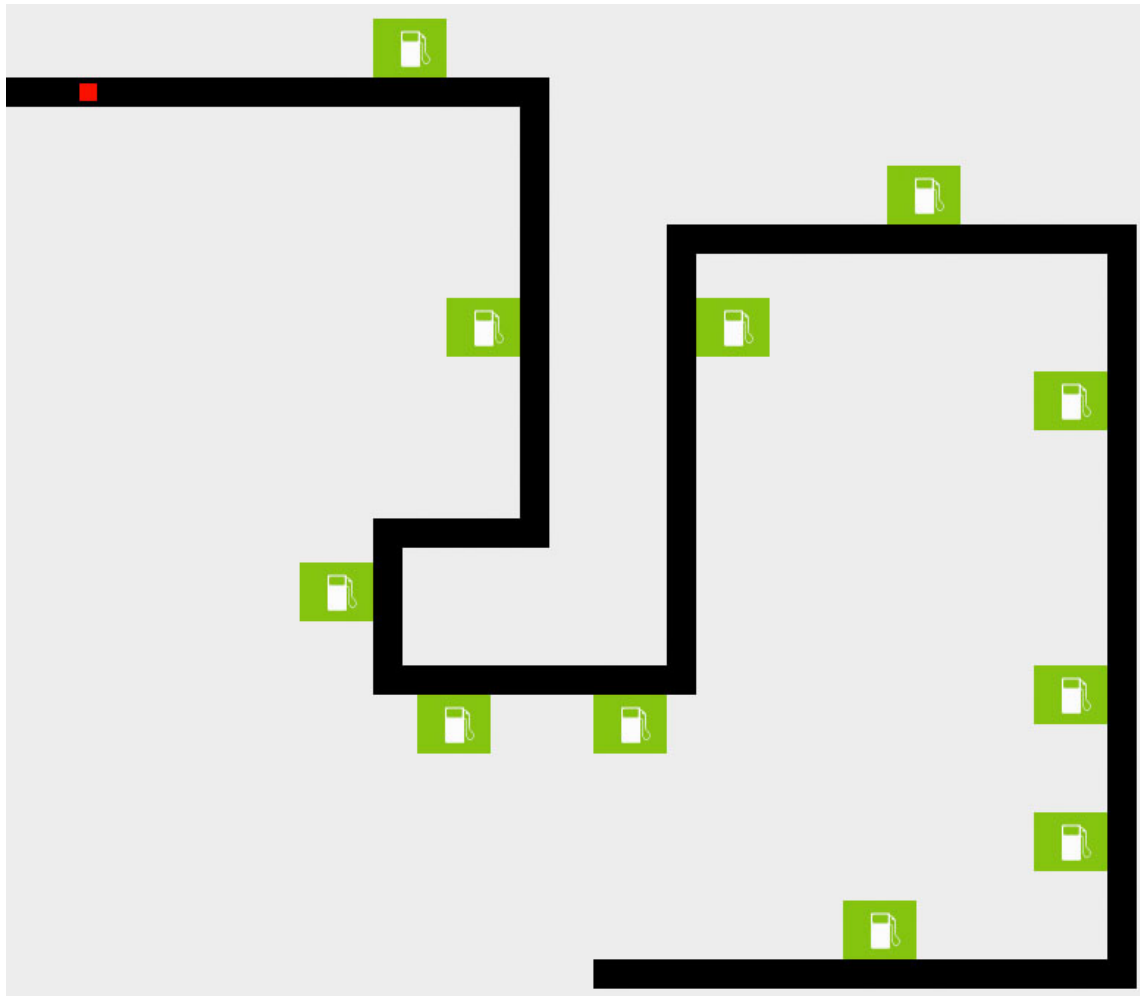


Figure 5.2: Locations of the fuel stations

In Figure-5.2, the locations of all the fuel stations are visible. However, from pilot testing of the interface, it was found that displaying of all the fuel stations on the display was not a good design. The interface was modified such that only the current and the next fuel stations are displayed on the map, as shown in Figure-5.3.

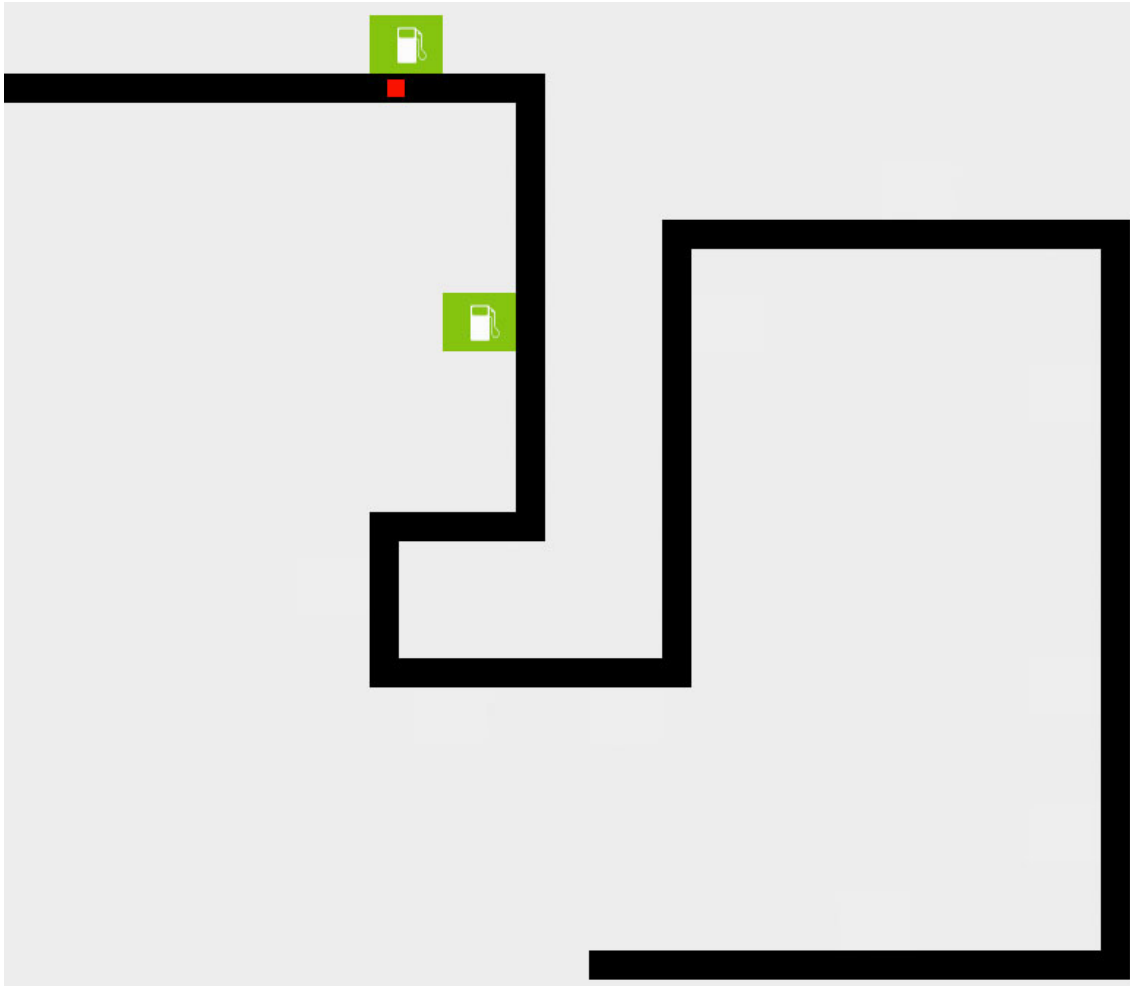


Figure 5.3: Final display of the component A

5.2.1.2 Component B

This area/component displays information about the upcoming fuel stations. When the vehicle is approaching a fuel station, the distance to the next two closest fuel stations are displayed as shown in Figure-5.4. The approximate fuel remaining in the vehicle is also displayed as a percentage value.

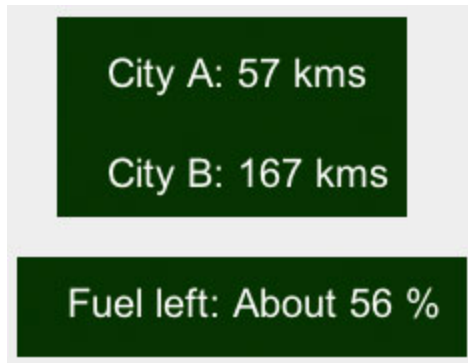


Figure 5.4: Display of component B

5.2.1.3 Component C

This area/component houses a button with the text “Refuel car”. This acts as a control for the user to fuel the car at a particular fuel station. The user clicks this button to refuel the car. This control is active only when the user is at any one of the fuel stations on the route. The control’s active and inactive state is shown in Figure-5.5.

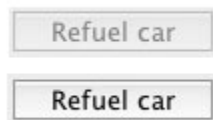


Figure 5.5: Inactive and active states of the control button in Component C

5.2.1.4 Component D

This area/component houses a low fuel icon as shown in Figure-5.6. This icon appears only when the fuel level of the vehicle is below 30%.



Figure 5.6: Component D – low fuel icon

5.2.1.5 Component E

This area consists of a slider, which gives a visual representation of the amount of fuel left in the vehicle, as shown in Figure-5.7.

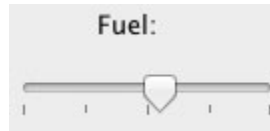


Figure 5.7: Component E – Fuel status of the vehicle

5.2.1.6 Component F

This component houses a text area that displays different text messages at various fuel levels as shown in Table-5.1.

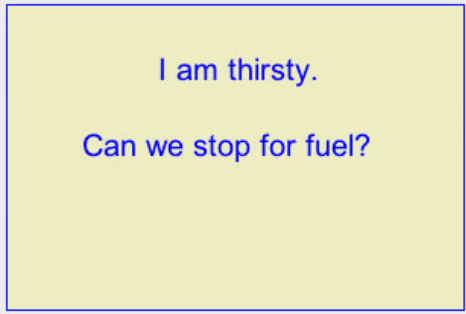

Fuel range	Text warning	Display as seen by the user
Between 60% and 75%	“I am thirsty. Can we stop for fuel?”	
Less than 45%	“AAARGH!! I am very thirsty. I need fuel!!!”	

Table 5.1: Messages displayed in Component F

5.2.1.7 Component G

This is the auditory component of the test-bed. There are three different variations in the use of this component. The variations are in terms of the messages issued. The following is the list of all the messages used in the test-bed

- 1) Beep (B) message: This is a beep sound that lasts for about two seconds.
- 2) Little thirsty (LT) message: This is an audio of the message “I am a little thirsty”.
- 3) Very thirsty (VT) message: This is an audio of the message “I am very thirsty”.
- 4) 75% fuel (75F) message: This is an audio of the message “75% fuel left”.
- 5) 45% fuel message: This is an audio of the message “45% fuel left”.

The variations of this component are shown in Table-5.2. The messages are played when the fuel level is at the indicated levels. They are played only once. For example, the 75% fuel message is issued at 75% fuel level of the vehicle. If the user does not drive the car at that moment, the fuel level remains to be 75%. However, the message is not issued another time.

Variation	Fuel level	Message
G1	30%	Beep message
G2	30%	Beep message
	45%	Little thirsty message
	75%	Very thirsty message
G3	30%	Beep message
	45%	45% fuel message
	75%	75% fuel message

Table 5.2: Details of Component G

5.2.2 Versions of the interface

Using the aforementioned components, five versions of the interface were developed as shown in Table-5.3.

Component	Version				
	T	EC1	EC2	AC2	AC2
A	Present	Present	Present	Present	Present
B	Present	Present	Present	Present	Present
C	Present	Present	Present	Present	Present
D	Present	Present	Present	Present	Present
E	Present	Present	Present	Absent	Absent
F	Absent	Present	Absent	Present	Absent
G	G1	G2	G3	G2	G3

Table 5.3: Details of the five versions of the interface

5.2.2.1 Version 1 – Traditional (T)

In this version, information about the fuel level is communicated to the user using the component D and E, i.e., the low fuel icon component and the fuel slider. When the fuel level is below 30%, the beep message is issued via the component G. Figure-5.8 shows a screenshot of the displayable components of this version.

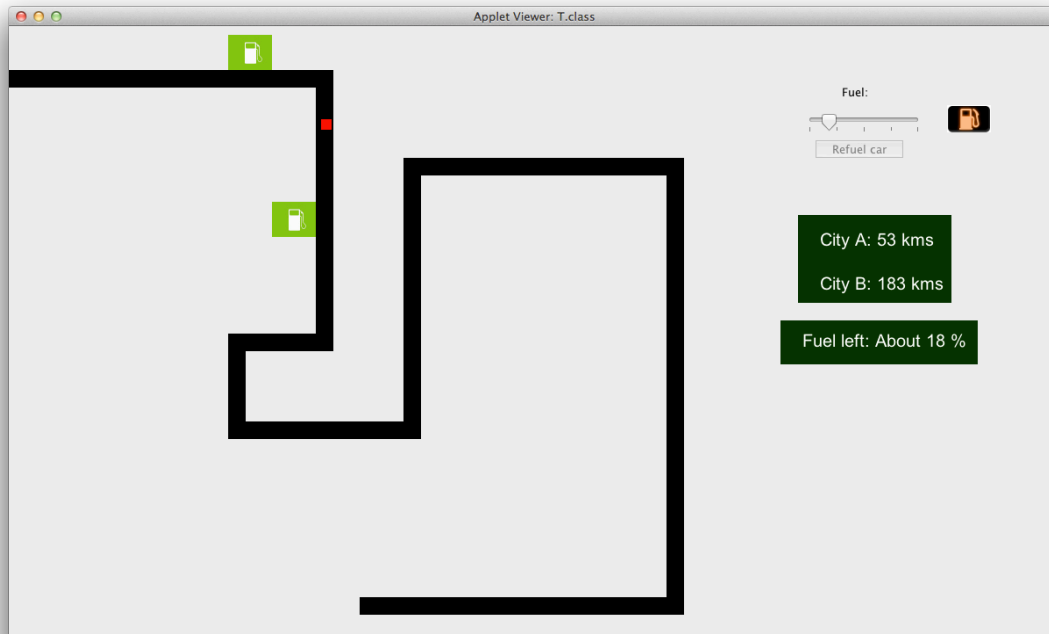


Figure 5.8: Version T

5.2.2.2 Version 2 – Eliminated Candidate 1 (EC1)

In this version, the fuel bar is eliminated from the interface. Instead, the information about the fuel level is communicated to the user using the components D, F and G, i.e., the low fuel icon, the text and the audio messages. The need for fuel level is represented as the vehicle being thirsty. The G2 variation of the audio component is used in this interface. Figure-5.9 shows a screenshot of the displayable components of this version.

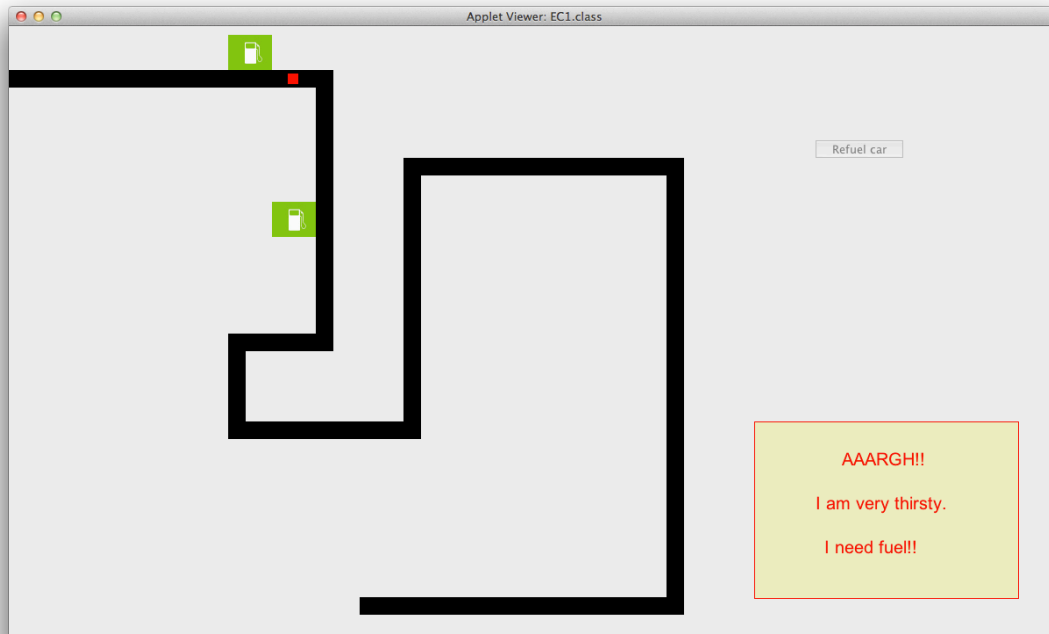


Figure 5.9: Version EC1

5.2.2.3 Version 3 – Eliminated Candidate 2 (EC2)

Similar to the above version, in this version, the fuel bar is also eliminated from the interface. Instead, the information about the fuel level is communicated to the user using the components D and G, i.e., the low fuel icon and audio messages. In the audio messages, the fuel level is directly announced to the user using the G3 variation of the audio component.

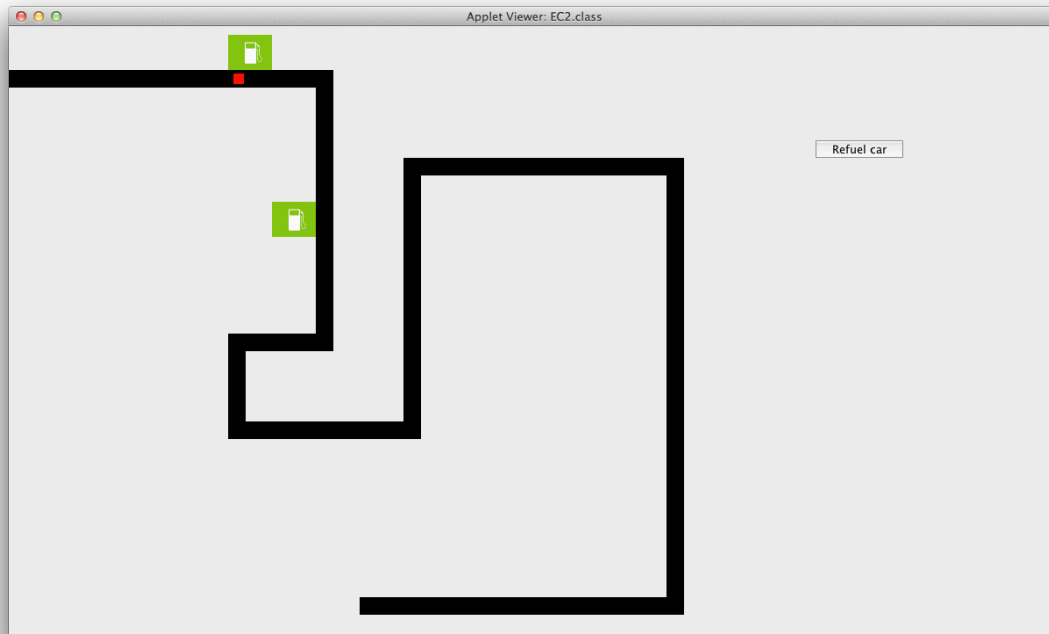


Figure 5.10: Version EC2

5.2.2.4 Version 4 – Augmented Candidate 1 (AC1)

In this version, the display of the information of the fuel status using the fuel meter (component E) and the low fuel icon (component D) is augmented by the text (component F) and audio (component G) messages. The need for fuel level is represented as the vehicle being thirsty. The G2 variation of the audio component is used in this interface. Figure-5.11 shows a screenshot of the displayable components of this version.

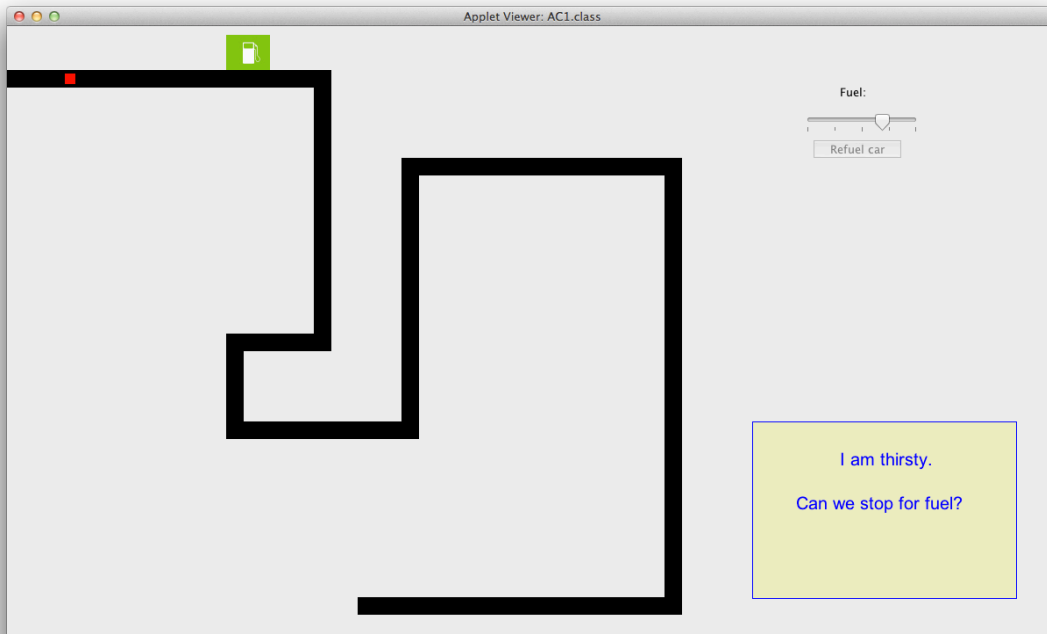


Figure 5.11: Version AC1

5.2.2.5 Version 5 – Augmented Candidate 2 (AC2)

In this version, the display of the information of the fuel status using the fuel meter (component E) and the low fuel icon (component D) is augmented by the audio messages (component G). The G3 variation of the audio component is used in this interface. The displayable components of this version looks same as Figure-5.8.

5.2.3 Additional information about the game

5.2.3.1 Task

In this game, the user is assigned the task to drive the vehicle from the start to the end of the path without running out of fuel. To achieve this task, the user has to monitor the fuel level and stop at various fuel stations to refuel the vehicle. If it happens that the

vehicle runs out of fuel, the user is notified about this and the car is automatically fuelled to the maximum level. Such instances are also recorded.

5.2.3.2 Measures

Several measures are recorded from the interface automatically. These measures are representative proxies of the user's action of fuelling the car at various fuel stations in the task.

- 1) Time spent refueling
- 2) Number of times fuel was finished
- 3) Minimum fuel level
- 4) Average fuel level
- 5) Minimum fuel level at refueling stations
- 6) Maximum fuel level at refueling stations
- 7) Average fuel level at refueling stations
- 8) Number of times refueled
- 9) Ending fuel level

5.3 Experimental design

It is required to see if there is any significant difference in the aforementioned measures due to various representations of the information about the machine's status (fuel).

There are five treatments in the experiment: T, EC1, EC2, AC1, and AC2. The participants of the experiment were asked to play the game on each of the five versions.

To eliminate effects due to the arrangement of the treatments, the participants are randomly divided into four blocks of treatments as shown in Table-5.4.

Block 1:	T	EC1	EC2	AC1	AC2
Block 2:	T	EC2	EC1	AC2	AC1
Block 3:	T	AC1	AC2	EC1	EC2
Block 4:	T	AC2	AC1	EC2	EC1

Table 5.4: Blocks with various treatments

5.3.1 Participants

A total of 12 participants (6 male, 6 female) were randomly selected from the students of University of Saskatchewan. The subjects were 21 to 29 years old, with a mean age of 25.16 years. All subjects were healthy in both their physical and mental states. The Ethics Committee in the University of Saskatchewan has approved this study (Appendix A).

5.3.2 Experimental Procedure

In one session lasting for about 45 minutes, the participant plays the game on each of the versions. The block assigned to the participant determines the order of the treatments. A small break of 5 minutes is given between two versions of the game play so as to forget the map/locations of the fuel stations etc. Prior to the start of the session, the participants were briefed upon the design of the interface, including the timing and the content of the warnings in all the interfaces.

5.4 Results

Paired sample student's t-test analyses were conducted on the data to investigate where there are any significant differences among the following pairs of interfaces:

- 1) T – AC1
- 2) T – AC2
- 3) T – EC1
- 4) T – EC2
- 5) AC1 – EC1
- 6) AC2 – EC2
- 7) AC1 – AC2
- 8) EC1 – EC2

Among the nine measures collected, five measures (Minimum fuel level, Average fuel level, Minimum fuel level at refueling stations, Maximum fuel level at refueling stations, Average fuel level at refueling stations) were based on the same underlying concept, i.e., the trends of the fuel level from the start to the end of the task. By running multiple t-tests, there is a risk of obtaining a false positive result. To address this issue, the Bonferroni method was used to modify the cut-off for statistical significance of the tests using the aforementioned measures. The cut-off for statistical significance used for these tests was $0.05/5 = 0.01$. For all other tests, a cut-off of 0.05 was used.

The following results were obtained after conducting a paired sample student's t-test on the data: There is a significant difference between the following measures in the versions EC1 and EC2 (see Table-5.5): Average fuel ($p=0.009$) and Average fuel at

refueling stations ($p=0.001$). These measures for the 12 participants are plotted in Figure-5.12 and Figure-5.13 respectively.

Measure	Mean	Std. deviation	Std. error mean	95% confidence interval		T	df	Sig. (2-tailed)
				Lower	Upper			
Average fuel	23.37	25.52	7.37	7.16	39.59	3.17	11	0.009
Average fuel at refuelling stations	45.32	34.99	10.10	23.09	67.56	4.49	11	0.001

Table 5.5: Paired differences between the pair EC1 and EC2.

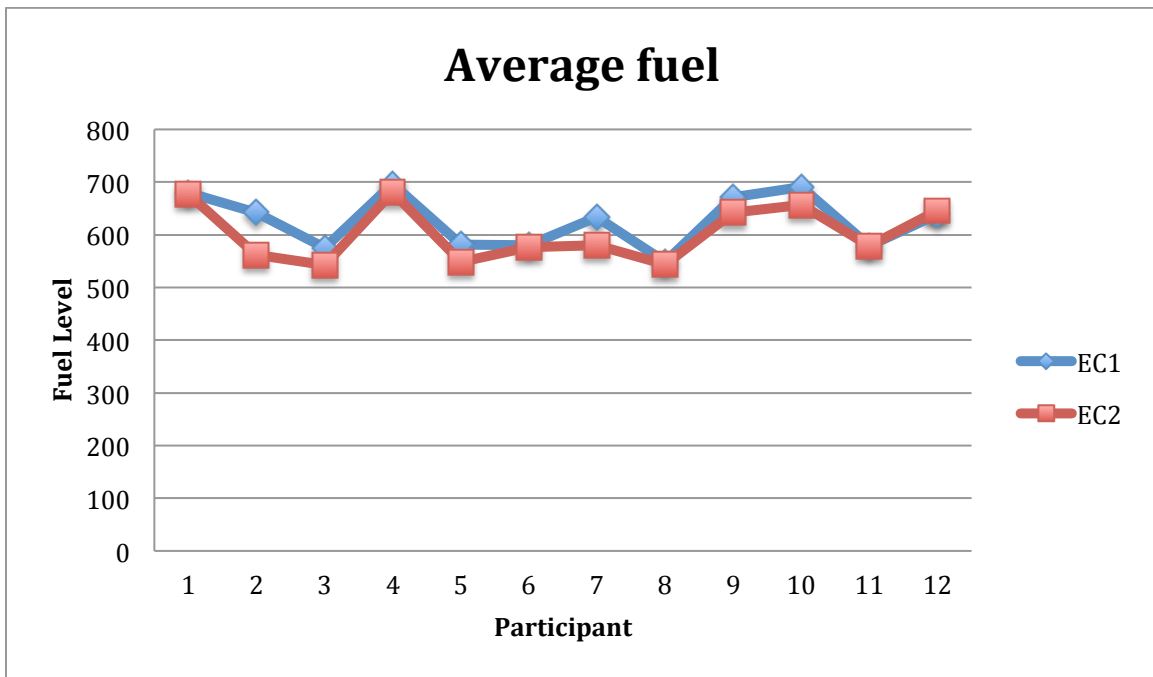


Figure 5.12: Average fuel level of the vehicle in EC1 and EC2

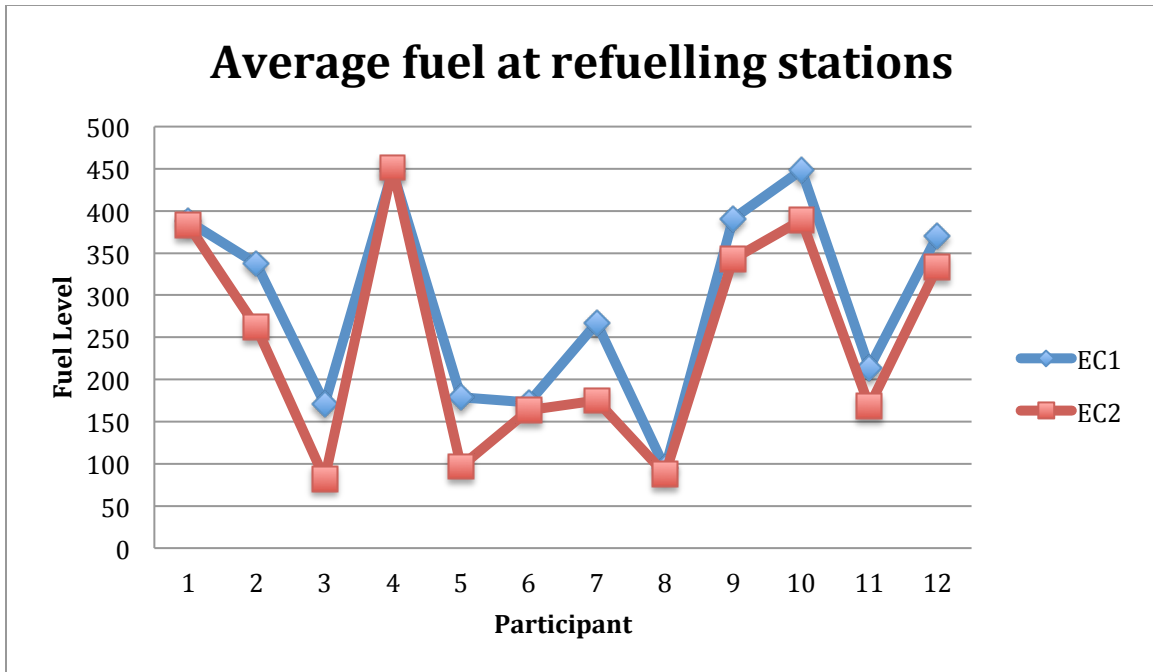


Figure 5.13: Average fuel level of the vehicle at refuelling stations in EC1 and EC2

It is evident from Figure-5.12 and Figure-5.13 that the scores of the measures for the version EC1 are consistently higher than those of EC2. The users refueled the car at higher fuel levels in EC1 when compared to EC2. It is recalled here that in both of these versions, the fuel meter is eliminated from the interfaces. The two interfaces differ in the degree of anthropomorphism designed in them. In EC1, the need for fuel level is communicated as the car being thirsty. The elements such as use of words and the medium of speech make the both interfaces anthropomorphic to some extent. However, the degree of anthropomorphism in EC1 is higher because it also uses the label of ‘thirst’. In EC2, no such ‘human-like’ label is used. Another difference between the two versions is that in EC1, there is a visual channel of warning, which is absent in EC2. The statistical difference between the measures of EC1 and EC2 can be accounted for by the above highlighted differences among the interfaces.

It was expected that there would be significant differences between the values of the measures between the augmented and eliminated interfaces. The belief was that due to the lack of a traditional visual element (i.e., the fuel meter), participants would be more cautious in the interfaces where the fuel meter was eliminated. However, the statistical analysis does not support this hypothesis for all cases. This hypothesis was found to be true only among the interfaces EC2 and AC2 in two measures (see Table-5.6): Time spent refueling ($p=0.032$) and Number of times fuel was finished ($p=0.039$). These measures for the 12 participants are plotted in Figure-5.14 and Figure-5.15 respectively. However, these statistics might have arisen due to the design of the game mechanics. In this game, it was implemented that when the vehicle ran out of fuel, it was automatically refueled. However, this action was hypothesized to take more time than refueling at a fuel station. This is because in real life scenarios, when a vehicle completely loses fuel, it takes quite some time for the driver to find some help and to be mobile again. Hence, in this version of the game, a small 'penalty' was to the total time spent refueling the car when such scenarios arise. The significant difference between the scores in both the interfaces can be easily explained due to this feature of the game. It is evident that the value of this 'penalty' also affects the p-value. For the above reasons, this finding does not have pragmatic meaning.

Measure	Mean	Std. deviation	Std. error mean	95% confidence interval		T	df	Sig. (2-tailed)
				Lower	Upper			
Time spent refuelling	837.25	1183.49	341.64	85.3	1589.2	2.45	11	0.032
Number of times fuel was finished	0.333	0.492	0.142	0.02	0.646	2.35	11	0.039

Table 5.6: Paired differences between the pair EC2 and AC2

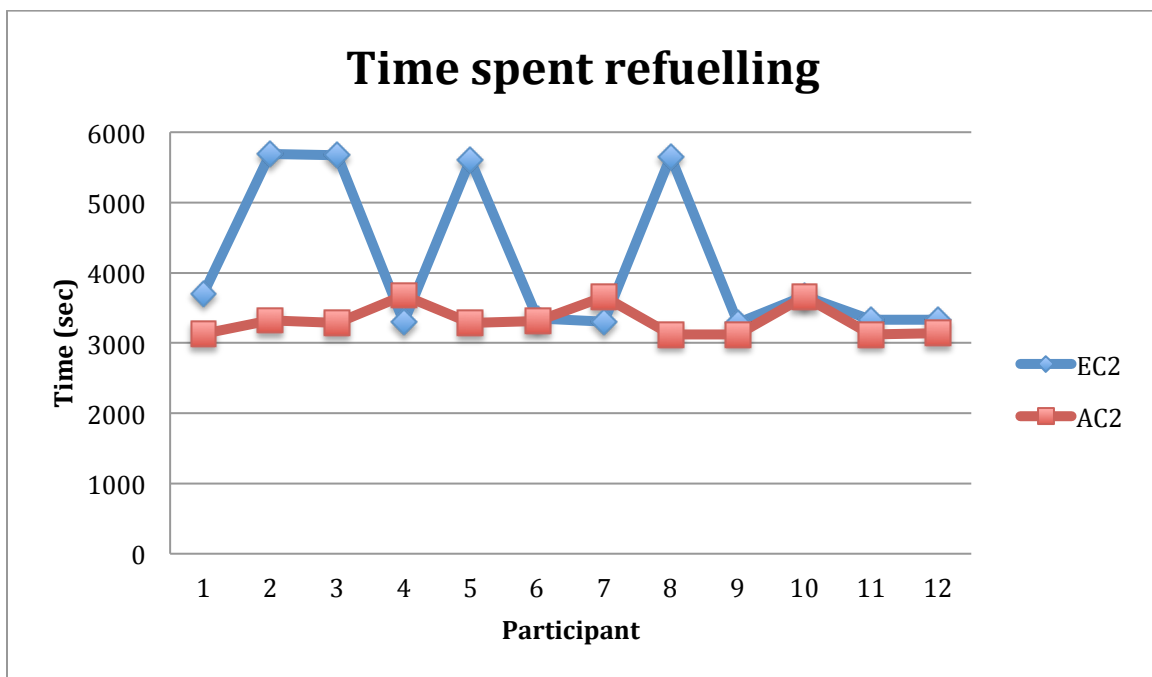


Figure 5.14: Total time spent refueling the vehicle during the task in EC2 and AC2

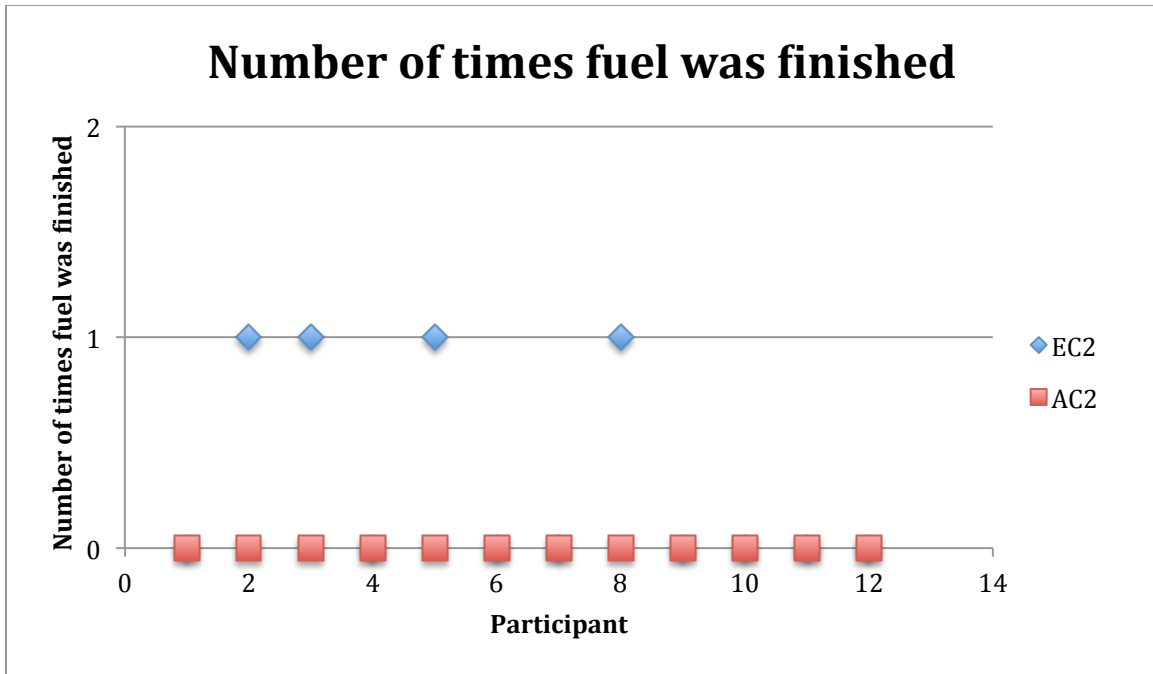


Figure 5.15: Number of times fuel was finished in EC2 and AC2

It was expected that there would be significant differences between the values of the measures between the traditional interface and any other interface. However, the statistical result does not support this belief. The hypothesis is recalled here: there is a significant effect on user's actions in a task when the machine's state is represented in a natural manner. A complete 'yes' answer could not be obtained in accordance with this hypothesis. This would have been achieved had there been any significant difference between the values of the measures of the traditional interface and any other interface. However, the difference in the values of the measures between the two natural representations of the machine state (EC1 and EC2) provides a partial 'yes' answer in support of the hypothesis. EC1 is more natural than EC2 due to the increased level of anthropomorphism, which is essentially the definition of the 'naturalness' in this study. The statistical evidence shows that there is significant difference in the measures between

the two interfaces, which is in support of the hypothesis. However, this hypothesis can only be grounded if differences could be found when comparing the interfaces individually to the traditional interface. The experimental setup was not very realistic to a human-vehicle-environment scenario, which might have been a cause of not achieving the above said statistical significance.

The findings in this thesis conclude that there is some evidence of an effect on the user's action in performing a task in a human-machine environment when the machine state is expressed in a natural way.

After the completion of the experiment, the participants were asked about their experience with the five versions of the game interface. This was done to obtain more insights in this research study by means of the technique of experience sampling.

About 83% of the participants preferred augmented versions of the interface to the traditional version. None of them preferred to have the versions in which the fuel meter slider is eliminated. Both the augmented candidates were preferred by an equal number of participants (Figure-5.16). Amongst the two versions of the auditory messages, i.e., "I am a little/very thirsty" and "75/45% fuel left", about 70% of the participants preferred the former one (Only 10 of the 12 participants answered this question with either of these choices).

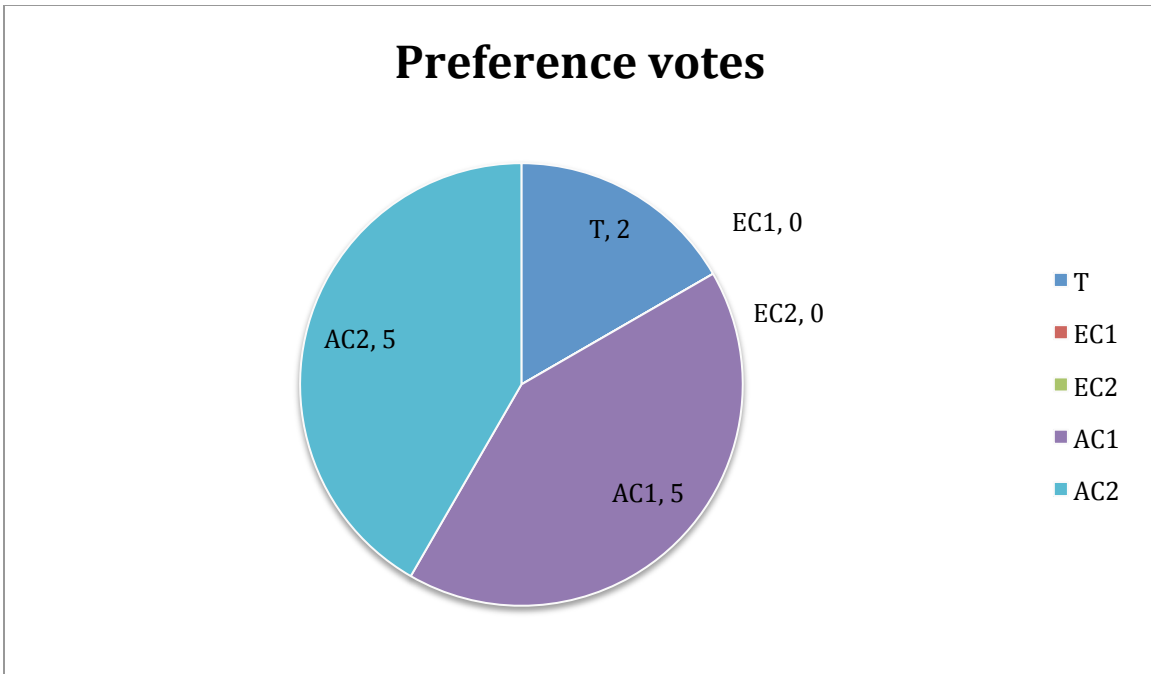


Figure 5.16: Preference of the interfaces amongst all interfaces

When polled amongst the clusters (a) T, EC1, AC1; and (b) T, EC2, AC2, most participants preferred the augmented versions to the traditional versions (Figure-5.17).

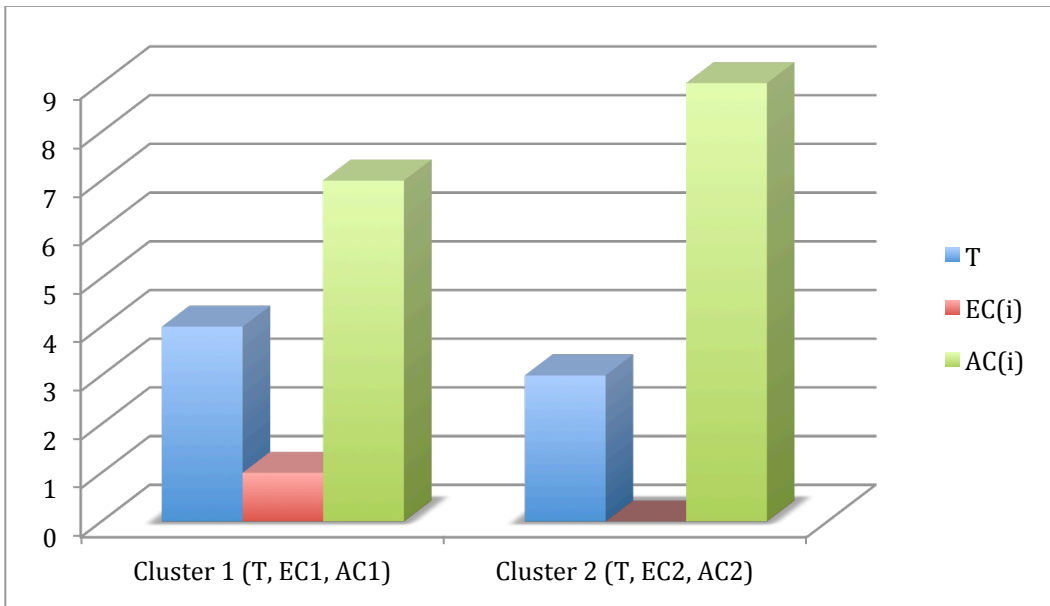


Figure 5.17: Preference of the interfaces in two clusters of polls

About 70% of the participants felt that there was an impact on their decisions of refueling due to the changes in the information presentation between the interfaces T and AC1. However, this number fell to 60% when answering the same question among the interfaces T and AC2. Only 36.4% of the respondents felt that the AC1 had greater impact on their refueling actions than AC2.

5.5 Conclusion

Representation of the machine's state in a natural manner is a new paradigm being explored in HMI. This study is one of the first of its kind to provide statistical evidence that there can be an effect on the user's actions when the machine's state is represented in a natural way.

In this thesis, the simulated test-bed of the human-machine environment was developed using a car game as an example. The results obtained shows good evidence that there is an effect on the user's actions in the task due to representation of the machine's state. Further studies are necessary to explore this paradigm using various human-machine interaction scenarios. Upon the completion of such studies, future designers can use this finding to develop better interaction amongst humans and machines.

Chapter 6: Conclusion

6.1 Overview

This thesis aims to improve the interaction between humans and machines. Three issues were identified and studied within the scope of this work. The first issue was related to the core capabilities that are necessary of a human-assistance system (HAS). In Chapter-3, a case study was used to demonstrate the incorporation of the core capabilities in the design of a HAS in a typical Computer Aided Design (CAD) scenario.

The second issue was related to improving the accuracy of human mind state inference. Instead of development of any new algorithm for improved accuracy, the approach taken was to develop a framework to integrate existing algorithms. The detailed description and discussion upon the results were discussed in Chapter-4.

The third issue was related to studying the effect of representation of the machine's state on the actions of the user who is interacting with the machine. It was hypothesized that there can be a possibility of a significant effect on the user's actions if the machine could communicate about its state in a natural manner. This hypothesis was tested by carrying experiments on a simulated test-bed of a driving game. The detailed description of the test-bed, experiment design and results was presented in Chapter-5.

6.2 Conclusions

The major conclusions obtained throughout this thesis are presented as follows:

- 1) During the development phase of Human Assistance Systems (HAS), designers of such systems should keep the following capabilities in consideration: task-

sharing, communication, transparency, rationale and cognition. These technologies can be implemented with the existing technologies including Fuzzy logic, Petri nets and ACT-R (Adaptive Character of Thought - Rational).

- 2) It is promising to apply the technique of expert opinion elicitation for human state inference. This study has demonstrated that a framework integrating opinions from various experts or algorithms to estimate the human fatigue state can yield an improved accuracy in inference. This conclusion, derived from the study of fatigue inference, is quite likely to be extended to other human states, for example emotion inference.
- 3) There is a great potential to affect the user's actions in a human-machine environment using a natural way of communication of the machine's state to the human. This thesis demonstrated this phenomenon with a simulated test-bed that was also developed in this thesis.

6.3 Contributions

The major contribution of this thesis lies in the improvement of interaction between humans and machines. The specific contributions are as follows:

- 1) This study has provided a platform for systematic design of Human Assistance Systems (HAS) from the principles of an ideal system for better interaction between humans and machines. The design of the HAS for CAD applications is the first of its kind from this standpoint.
- 2) This study has provided a general framework for integrating opinions from several experts for improved accuracy of estimation. The implementation of the framework for fatigue inference is the first of its kind to be reported. The

framework can be extended to suit other applications in inference such as emotion inference (e.g., anxiety, frustration, etc.)

- 3) This study has generated first hand evidence of the effect of the representation of the machine's state on a user's actions in a human-machine environment. This is one of the required groundwork to be conducted for evolution of human-machine interactions.

6.4 Limitations and future work

The specific limitations of the studies of this thesis were discussed in their respective chapters. This section discusses the general limitations and future direction of research that can be taken to extend the current work.

- 1) In the context of HAS, the proposed list of core capabilities identified is not exhaustive of all scenarios. Some human-machine interaction scenarios require even more capabilities or constraints – for example, consider a case of a human-vehicle situation; one additional core capability of the machine could be to ensure safety of the passengers in the vehicle. The technology behind automatic braking can then be viewed as the human assistance system, which has to then be critically analyzed from the viewpoint of the list of core competencies, identified in this thesis. Indeed, this thesis provided the first study of HAS to start such a list. However, several other cases spanning across varied human-machine interaction scenarios have to be studied to populate the list of core capabilities. This is a challenging task and will demand collaboration between several researchers in a multi-disciplinary fashion.

- 2) In the context of human-state inference, there is an immediate need for developing a standard for human-state representation. Currently, the representation of the human state is so varied that one cannot easily compare and contrast between the results of two techniques of inferring the same human state.
- 3) In the context of human-machine communication, this thesis studied the effect of the machine's state represented in only one simulated scenario, i.e., a driving task. Following up on the potential of such a representation, there is a need for conducting user studies in various human-machine systems. The paradigm of representation of the machine's state in a natural manner has to be thoroughly studied before they can be used in real-life interactions with machines.

References

- Anderson, J. R. and Lebiere, C. (1998), *The atomic components of thought*, 1998, Mahwah: NJ: Lawrence Erlbaum Associates.
- Anderson, J. R., Matessa, M. and Douglass, S. (1995), *The ACT-R theory and visual attention*. In: *Proc. Seventeenth Annual Conf. of the Cognitive Science Society*, University of Pittsburg. 22-25 July 1995.
- Anderson, J. R., Matessa, M. and Lebiere, C. (1997), *ACT-R: A theory of higher level cognition and its relation to visual attention*, *Human-Computer Interaction*, 1997, 12(4), 439-462.
- Bartneck C. (2001), "How convincing is Mr. Data's smile: Affective expressions of Machines", *User Modeling and User-Adapted Interaction*, 11, 279-295.
- Breazeal, C. (2002), *Designing Sociable Robots*. The MIT Press.
- Brown D. R. and Hwang K. Y. (1993), *Solving fixed configuration problems with genetic search*, *Research in Engineering Design*, 1993, 5(2), 80-87.
- Cai, H. (2011), "Fine-tuning human-machine cooperation: cognitive processing oriented assistance coordination", Ph.D Thesis, *Industrial Engineering Dissertations*, Northeastern University <http://hdl.handle.net/2047/d20000969>
- Campbell M. I. (2000), "The A-Design invention machine: a means of automating and investigating conceptual design". Ph.D Thesis, 2000, Carnegie Mellon University, Pittsburgh, PA.

- Campbell M. I., Cagan J. and Kotovsky K. (1999), A-design: an agent based approach to conceptual design in a dynamic environment, *Research in Engineering Design*, 1999, 11(3), 172-192.
- Campbell M. I., Cagan J. and Kotovsky K. (2000), Agent-based synthesis of electromechanical design configurations, *Journal of Mechanical Design*, 2000, 122(1), 61-69.
- Campbell M. I., Cagan J. and Kotovsky K. (2003), The A-Design approach to managing automated design synthesis, *Research in Engineering Design*, 2003, 14(1), 12-24.
- Carlson-Skalak S., White M. D. and Teng Y. (1998), Using an evolutionary algorithm to catalog design, *Research in Engineering Design*, 1998, 10(2), 63-83.
- Clemen, R. T. and Winkler, R. L. (1999), "Combining probability distribution from experts in risk analysis," *Risk Analysis*, vol. 19, pp. 187-203, 1999
- Deb, K. (2001), *Multiobjective Optimization Using Evolutionary Algorithms*, Wiley, 2001.
- Franklin S. and Graesser A. (1997), Is it an agent, or just a program?, in J. P. Muller, M. Wooldridge, and N. R. Jennings (Eds.), *Intelligent Agents III (LNAI Volume 1193)*, Springer-Verlag: Berlin, Germany, 1997, pp. 21–36.
- Gadando S. C. and Hallam J. (1998), "Exploring the role of emotions in autonomous robot learning", *AAAI Technical report FS-98-03*, 84-89.
- Goleman D. (1995), *Emotional Intelligence*. New York: Bantam Books, 1995
- Gray W. D. (2007), *Integrated Models of Cognitive Systems*, Oxford University Press, USA, 2007

- Hamonda G. and Saccomanno F. F. (1995), "Neural network model for truck driver fatigue accident detection," in Proc. IEEE Canadian Conference on Electrical and Computer Engineering, Montreal, Canada, vol. 1, 1995, pp. 362-365.
- He C. D. and Zhao C. C. (1993), "Evaluation of the critical value of driving fatigue based on the fuzzy sets theory," Environmental Research, vol. 61, pp. 150-156, 1993.
- Huber, P., Jensen, K. and Shapiro, R. M. (1990) Hierarchies in coloured Petri nets, In: G. Rozenberg (ed.): Advances in Petri Nets, 1990, Lecture Notes in Computer Science, Springer-Verlag, 1990, pp. 313-341.
- Horng W., Chen C., Chang Y. and Fan C. (2004), Driver fatigue detection based on the eye tracking and dynamic template matching, in Proc. IEEE International Conference on Networking, Sensing and Control, Taiwan, vol. 1, 2004, pp. 7-12.
- Ji Q., Lan P. and Looney C. (2006), "A probabilistic framework for modeling and real-time monitoring human fatigue", IEEE Transactions on Systems Man, and Cybernetics: Part A, vol. 36, pp. 862-875, 2006.
- Ji Q., Zhu Z. and Lan P. (2004), "Real-time nonintrusive monitoring and prediction of driver fatigue", IEEE Transactions on Vehicular Technology, vol. 53, pp. 1052-1068, 2004.
- Jung T. P, Makeig S., Stensmo M. and Sejnowski T.J. (1997), "Estimating alertness from the EEG power spectrum", IEEE Transactions on Biomedical Engineering, vol. 44, pp. 60-69, 1997.

- Kusiak A., Szcherbick E. and Vujosevik R. (1991), Intelligent design synthesis: an object oriented approach, *International Journal of Production Research*, 1991, 29(7), 1291-1308.
- Lang W. Y. and O’Gardy P. (2001), OODS: an objected-oriented design system - an example of electronic assemblies, *Computers and Industrial Engineering*, 2001, 40(3), 229-247.
- Lebiere, C. and Anderson, J. R. (1993), A connectionist implementation of the ACT-R production system, In *Proc. Fifteenth Annual Conference of the Cognitive Science Society*, University of Colorado-Boulder, 18-21 July 1993.
- Lee-Johnson, C.P. and Carnegie, D.A. (2010), "Mobile Robot Navigation Modulated by Artificial Emotions," *Systems, Man, and Cybernetics, Part B: Cybernetics*, *IEEE Transactions on Systems, Man and Cybernetics - Part B: Cybernetics*, vol.40, no.2, pp.469-480
- Li X. and Ji Q. (2005), “Active affective state detection and user assistance with dynamic Bayesian networks,” *IEEE Transactions on Systems Man, and Cybernetics: Part A*, vol. 35, pp. 93-105, 2005.
- Lin, Y., H. Leng, G. Yang, H. Cai (2007) An Intelligent Noninvasive Sensor for Driver Pulse Measurement, *IEEE Sensors Journal*, Vol. 7, No. 5, pp. 790-799.
- Lin, Y. (2010), “Affective Driving”, Chapter 14. ed. by Shuichi Fukuda, *Emotional Engineering: Service Development*, Springer , 2010.
- Liu Y., Feyen R. and Tsimhoni O. (2006), Queueing Network-Model Human Processor (QN-MHP): a computational architecture for multitask performance in human-

- machine systems, *ACM Transactions on Computer-Human Interaction*, 2006, 13(1), 37-70.
- Mandryk R. and Atkins M. (2007), “A fuzzy physiological approach for continuously modeling emotion during interaction with play technologies”, *International Journal of Human-Computer Studies*, 65(4), pp. 329-347
- Moss. J., Cagan K. and Kotovsky K. (2004), Learning from design experience in an agent-based design system, *Research in Engineering Design*, 2004, 15(2), 77-92.
- Murata, T. (1989) Petri nets: properties, analysis and applications, *Proc. IEEE*, 1989, 77(4), 541-580.
- Nasoz, F., Alvarez, K., Lisetti, C. L., 2003, Emotion Recognition from Physiological Signals for User Modeling of Affect, *Proceedings UM’ 2003, 9th International Conference on User Model*, Pittsburg, pp. 1-9.
- Nasoz, F., Alvarez, K., Lisetti, C. L., 2004, Emotion recognition from physiological signals using wireless sensors for presence technologies, *Cognition, Technology and Work*, vol., 6, issue 1, pp. 4-14.
- Nass C., Jonsson I., Harris H., Reaves B., Endo J., Brave S. and Takayama L. (2005), “Improving automotive safety by pairing driver emotion and car voice emotion”, *CHI’05 extended abstracts on Human factors in computing systems*, pp. 1976
- Norman, D. A. (1992), “Turn Signals Are the Facial Expressions of Automobiles”, Addison Wesley, Online: http://www.jnd.org/dn.mss/chapter_11_turn_sign.html

- Norman, D. A. (2004), "Emotional Design: Why We Love (or Hate) Everyday Things", Basic Books.
- Norman D. A. (2007), "The Design of Future Things", New York: Basic Books.
- Pedrycz W, Gomide F. (1999), "Generalized Petri nets as pattern classifiers", Pattern Recognition Letters 1999; 29(14): 1489–98.
- Picard R. W., Vyzas E. and Healey J. (2001), "Towards machine emotional intelligence: analysis of affective physiological state", IEEE Transactions on Pattern Analysis and Machine Intelligence, 23(10), 1175-1191
- Potkonjak V., Radojicic J., Tzafestas S. and Kostic D., (2002a), "Modeling Robot "Psycho-Physical" State and Reactions – A New Option in Human–Robot Communication. Part 1: Concept and Background", Journal of Intelligent and Robotic Systems, 35, 339-352
- Potkonjak V., Radojicic J., Tzafestas S. and Kostic D., (2002b), "Modeling Robot "Psycho-Physical" State and Reactions – A New Option in Human–Robot Communication. Part 2: Modeling and Simulation", Journal of Intelligent and Robotic Systems, 35, 353-364
- Ruff, A. H., Narayanan, S. and Draper, M. H. (2002), Human interaction with levels of automation and decision-aid fidelity in the supervisory control of multiple simulated unmanned air vehicles, Presence, 2002, Vol. 11, 335-351.
- Russell, J. A., Weiss, A., & Mendelsohn, G. A. (1989), "Affect grid: A single-item scale of pleasure and arousal", Journal of Personality and Social Psychology, 57, 493–502.

- Saerbeck M. and Bartneck C. (2010), "Perception of affect elicited by robot motion", Proc. 5th ACM/IEEE Int. Conf. Human-robot interaction, pp. 53-60.
- Saerbeck M., Schut T., Bartneck C. and Janse M. D. (2010), "Expressive robots in education: Varying the degree of social supportive behavior of a robotic tutor", CHI 2010: Classroom technologies, pg. 1613-1622, Atlanta, GA, USA
- Saldien J, Goris K, Vanderborcht B, Vanderfaeillie J, Lefeber D (2010) Expressing emotions with the social robot Probo. Int J Soc Robot. Available online: doi: 10.1007/s12369-010-0067-6
- Sheridan T. B. and Parasuraman R. (2005), Human-automation interaction, Reviews of Human Factors and Ergonomics, 1, pp. 89-129
- Suzuki K., Camurri A., Ferrentino P. and Hashimoto S., (1998), "Intelligent agent system for human-robot interaction through artificial emotion", 1998 IEEE Conference on Systems, Man and Cybernetics, 2, 1055-1060
- Tay E. H., Flowers W. and Barrus J. (1998), Automated generation and analysis of dynamic system designs, Research in Engineering Design, 1998, 10(1), 15-29.
- Yager, R. R. (2004) "OWA aggregation over a continuous interval argument with applications to decision making", IEEE Transactions on Systems, Man, and Cybernetics: Part B, vol. 34, no. 5, pp. 1952-1963, 2004.
- Yang G., Lin Y. and Bhattacharya P. (2008), "A method for multi-modality inferring of human cognitive states based on the integration of neuro-fuzzy network and information fusion techniques," EURASIP Journal on Advances in Signal Processing, vol. 2008, Article ID 371621, 14 pages, 2008.

- Yang G., Lin Y. and Bhattacharya P. (2010), "A Driver Fatigue Recognition Model Based on Information Fusion and Dynamic Bayesian Network," *Information Sciences*, vol. 180, pp. 1942-1954, 2010.
- Wu, C., Tsimhoni, O., and Liu, Y. (2008), "Development of an adaptive workload management system with the Queuing Network-Model Human Processor (QN-MHP)", *IEEE Transactions on Intelligent Transportation Systems*, 9(3), pp. 463-475.
- Zadeh, L. A. (1975), *Fuzzy logic and approximate reasoning*, Synthese, 1975, 30, 407-428.
- Zhang, W.J., Ouyang, P.R., Sun, Z.H. (2010), "A novel hybridization design principle for intelligent mechatronics systems", *Proceedings of International Conference on Advanced Mechatronics (ICAM2010)*, 2010, Osaka, Japan, October 4-6.
- Zurawski, R. and Zhou, M. (1994), *Petri nets and industrial applications - a tutorial*, *IEEE Transactions on Industrial Electronics*, 1994, 41(6), 567-583.

Appendix A: Certificate of ethics approval for the experiment



UNIVERSITY OF
SASKATCHEWAN

Behavioural Research Ethics Board (Beh-REB)

Certificate of Approval

PRINCIPAL INVESTIGATOR
Chris Zhang

DEPARTMENT
Mechanical Engineering

BEH#
11-180

INSTITUTION(S) WHERE RESEARCH WILL BE CONDUCTED
University of Saskatchewan

STUDENT RESEARCHER(S)
Shrey Modi

FUNDER(S)
INTERNALLY FUNDED

TITLE
Effect of Interface Display of Vehicle's State on User Performance in a Car Game

ORIGINAL REVIEW DATE
08-Jul-2011

APPROVAL ON
26-Jul-2011

APPROVAL OF:
Ethics Application
Consent Protocol

EXPIRY DATE
25-Jul-2012

Full Board Meeting

Date of Full Board Meeting:

Delegated Review

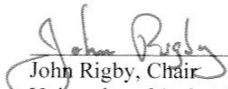
CERTIFICATION

The University of Saskatchewan Behavioural Research Ethics Board has reviewed the above-named research project. The proposal was found to be acceptable on ethical grounds. The principal investigator has the responsibility for any other administrative or regulatory approvals that may pertain to this research project, and for ensuring that the authorized research is carried out according to the conditions outlined in the original protocol submitted for ethics review. This Certificate of Approval is valid for the above time period provided there is no change in experimental protocol or consent process or documents.

Any significant changes to your proposed method, or your consent and recruitment procedures should be reported to the Chair for Research Ethics Board consideration in advance of its implementation.

ONGOING REVIEW REQUIREMENTS

In order to receive annual renewal, a status report must be submitted to the REB Chair for Board consideration within one month of the current expiry date each year the study remains open, and upon study completion. Please refer to the following website for further instructions: http://www.usask.ca/research/ethics_review/


John Rigby, Chair
University of Saskatchewan
Behavioural Research Ethics Board

Please send all correspondence to:

Research Ethics Office
University of Saskatchewan
Box 5000 RPO University, 1602-110 Gymnasium Place
Saskatoon SK S7N 4J8
Telephone: (306) 966-2975 Fax: (306) 966-2069

Appendix B: Data used for implementation of the framework for fatigue inference

Training data for inferring fatigue (y_1) from the sleeping quality (z_1):

z_1	y_1	z_1	y_1	z_1	y_1	z_1	y_1
0.05	0.95	0.12	0.9	0.93	0.08	0.5	0.5
0	1	1	0	0.85	0.15	0.45	0.55
0.09	0.93	0.9	0.1	0.13	0.85	0.54	0.47
0.1	0.9						

Training data for inferring fatigue (y_2) from driving hours (z_2):

z_2	y_2	z_2	y_2	z_2	y_2	z_2	y_2
0.0417	0.06	0.249	0.3	0.542	0.55	0.83	0.8
0.083	0.12	0.332	0.33	0.581	0.6	0.875	0.85
0.125	0.18	0.415	0.4	0.625	0.65	0.913	0.9
0.167	0.25	0.458	0.45	0.664	0.7	0.958	0.95
0.208	0.22	0.498	0.5	0.747	0.75	1	1

Training data for inferring fatigue value (y_3) from EEG analysis (z_3 and z_4):

z_3	z_4	y_3	z_3	z_4	y_3	z_3	z_4	y_3
0.1	0.2	0.01	0.5	0.5	0.5	0.93	0.95	0.91
0.4	0.2	0.01	0.49	0.49	0.49	0.9	0.89	0.91
0.3	0.1	0.01	0.55	0.45	0.49	0.85	0.87	0.91
0.2	0.4	0.01	0.47	0.47	0.49	0.8	0.75	0.9
0.01	0.01	0.01	0.45	0.5	0.49	0.75	0.8	0.9

Training data for inferring fatigue value (y_4) from ECG analysis (z_5):

z_5	y_4	z_5	y_4	z_5	y_4	z_5	y_4
0	0	0.41	0.25	0.62	0.53	0.9	0.91
0.083	0.09	0.45	0.48	0.667	0.55	0.917	0.92
0.12	0.1	0.47	0.5	0.75	0.91	0.95	0.92
0.167	0.15	0.49	0.5	0.79	0.91	1	0.93
0.2	0.18	0.5	0.5	0.81	0.91		
0.25	0.2	0.55	0.51	0.833	0.91		
0.333	0.22	0.583	0.53	0.85	0.91		

Training data for inferring fatigue value (y_5) from eye movement analysis (z_6):

z_6	y_5	z_6	y_5	z_6	y_5	z_6	y_5
0	0	0.367	0.13	0.617	0.25	0.917	0.9
0.083	0.1	0.417	0.13	0.667	0.26	0.92	0.91
0.117	0.1	0.45	0.15	0.75	0.87	0.95	0.92
0.167	0.1	0.5	0.16	0.783	0.87	0.97	0.93
0.25	0.11	0.55	0.2	0.833	0.88	1	0.93
0.333	0.12	0.583	0.2	0.867	0.88		

Testing data for inferring fatigue value (y) from the cues:

z_1	z_2	z_3	z_4	z_5	z_6	y
0.9	0.06	0.4	0.3	0.12	0.19	0.1
0.9	0.17	0.4	0.5	0.42	0.19	0.18
0.5	0.5	0.5	0.5	0.5	0.7	0.5
0.5	0.5	0.45	0.45	0.42	0.56	0.39
0.01	1	1	0.99	1	1	1
0.7	0.17	0.3	0.4	0.28	0.69	0.24
0.8	0.28	0.5	0.4	0.44	0.19	0.24
0.88	0.17	0.45	0.4	0.42	0.7	0.26
0.1	0.83	0.9	0.89	1	0.75	0.884
0.15	0.96	0.75	0.8	0.92	0.75	0.884
0.23	0.83	0.75	0.78	0.75	0.74	0.83
0.19	0.86	0.7	0.75	0.73	0.89	0.83
0.1	0.63	0.7	0.7	0.71	0.74	0.756
0.34	0.68	0.8	0.75	0.72	0.73	0.756
0.08	0.92	0.98	0.98	0.82	1	0.926
0.1	0.93	0.93	0.95	1	1	0.926
0.065	0.92	0.95	0.95	0.75	1	0.926

Appendix C: Information fusion using OWA technique

The information fusion principle for integrating the opinions from various experts using the OWA technique is described in this appendix (adapted from Yang et al. 2008).

Let y_i ($i = \{1 \dots p\}$) be the opinion from the i^{th} expert among a group of p experts denoted as $\mathbf{o} = [y_1, y_2, \dots, y_p]^T$. Let $\mathbf{w} = [w_1, w_2, \dots, w_p]^T$ denote the associated weight vector. Construct $\mathbf{b} = [b_1, b_2, \dots, b_p]^T$ such that b_i ($i = \{1 \dots p\}$) is the i^{th} largest element of \mathbf{o} . As per the OWA method, the integrated opinion y can be calculated as follows:

$$y = \mathbf{w}^T \mathbf{b} = \sum_{i=1}^p w_i b_i \quad (\text{C-1})$$

$$0 \leq w_i \leq 1 \quad i = 1, 2, \dots, p \quad (\text{C-2})$$

$$\sum_{i=1}^p w_i = 1 \quad (\text{C-3})$$

A number of techniques are available to determine the weight vector \mathbf{w} . In Yang et al. (2008), a combined technique is used as follows:

Let $\widehat{\mathbf{w}} = \{\widehat{w}_i (i = 1, 2, \dots, p)\}$ be the estimation of \mathbf{w} , and specify

$$\widehat{w}_i = \frac{e^{\lambda_i}}{\sum_{j=1}^p e^{\lambda_j}} \quad i = 1, 2, \dots, p \quad (\text{C-4})$$

In order to ensure the constraints of $0 \leq \widehat{w}_i \leq 1$ ($i = \{1 \dots p\}$) and $\sum_{i=1}^p \widehat{w}_i = 1$, λ_i is taken as the unknown parameter to be determined in the learning process. There are k outputs of the network, denoted by $\mathbf{o}_k = [y_{k1}, y_{k2}, \dots, y_{kp}]^T$ ($k = 1, 2, \dots, K$). According to OWA, \mathbf{o}_k is reordered to $\mathbf{b}_k = [b_{k1}, b_{k2}, \dots, b_{kp}]^T$, where b_{ki} is the i^{th} largest element of

the collection of \mathbf{o}_k . Let \widehat{y}_d^k be the current estimated aggregated values corresponding to \mathbf{b}_k and $\widehat{\mathbf{w}}$. Then \widehat{y}_d^k can be calculated by

$$\begin{aligned}\widehat{y}_d^k &= \widehat{\mathbf{w}}^T \mathbf{b}_k = \sum_{i=1}^p \widehat{w}_i b_{ki} \\ &= \frac{b_{k1} e^{\lambda_1}}{\sum_{j=1}^p e^{\lambda_j}} + \frac{b_{k2} e^{\lambda_2}}{\sum_{j=1}^p e^{\lambda_j}} + \dots + \frac{b_{kp} e^{\lambda_p}}{\sum_{j=1}^p e^{\lambda_j}}\end{aligned}\quad (\text{C-5})$$

Let y_d^k be the expected aggregated values corresponding to \mathbf{o}_k , then the error e_k between y_d^k and \widehat{y}_d^k can be calculated by

$$\begin{aligned}e_k &= \frac{1}{2} (\widehat{y}_d^k - y_d^k)^2 \\ &= \frac{1}{2} (\sum_{i=1}^p \widehat{w}_i b_{ki} - y_d^k)^2\end{aligned}\quad (\text{C-6})$$

Using the steepest gradient descent method, the parameters $\lambda_i (i = 1, 2, \dots, p)$ are updated with the following equation:

$$\lambda_i(k+1) = \lambda_i(k) - 2\beta w_i (b_{ki} - \widehat{y}_d^k) e_k \quad (\text{C-7})$$

where β is the learning rate. Consequently, parameters w_i are calculated at each iteration step for the current values of parameters $\lambda_i(k)$ ($i = 1, 2, \dots, p$).