Jarvis in the car: Report on characterizing and designing in-vehicle intelligent agents workshop

Manhua Wang[1], Phillip Hock[2], Seul Chan Lee[3], Martin Baumann[2], and Myounghoon Jeon[1]
[1] Virginia Tech, Blacksburg, United States
[2] Ulm University, Ulm, Germany
[3] Gyeongsang National University, Jinju, Republic of Korea

As intelligent agents have become more popular at home, they have been progressively introduced into driving environments. Although previous research has discussed agent features and their effects on driver perception and performance, attributes that define in-vehicle agents and distinguish them from other intelligent agents have not been discussed clearly. Thus, we organized a workshop on characterizing and designing in-vehicle intelligent agents at the 13th International Conference on Automotive User Interfaces (AutoUI 2021). In this report, we integrated ideas generated during the workshop and identified user-centered action and autonomy as two attributes that define an agent, with functions and features as specific characteristics that vary agent design. The outcomes of this workshop can facilitate in-vehicle intelligent agent design and deliver optimal user experience, while providing insights on manipulating variables in controlled studies.

INTRODUCTION

Intelligent agents (IAs), which perform user-given tasks autonomously and intelligently (March et al., 2000), have been introduced into our daily lives as voice assistants. IAs have gradually penetrated the automobile industry due to their advantages of not competing with driving, which is a visually demanding task. With the emerging technologies that advance vehicular automation and driver assistant systems, in-vehicle intelligent agents (IVIAs) no longer only serve as a voice control system that is connected through smartphones via Bluetooth or cable. IVIAs are also expected to level up a driver assistant system to communicate system boundaries and secure safe driving, especially under semi-automated driving conditions where drivers and vehicles share the driving responsibility.

While previous research has provided valuable insights on the features of IVIAs and the resulting user perception, characteristics that uniquely scope IVIAs and distinguish them from at-home voice assistants remain untouched. To further characterize and optimize IAs situated in the driving contexts, we hosted our first workshop on characterizing and designing in-vehicle intelligent agents at the 13th International ACM Conference on Automotive User Interfaces (AutoUI 2021) (Wang, Hock, et al., 2021). This two-day workshop aimed to gather the opinions of the experts and practitioners in the field of automated vehicles to scope the definition of agents, and then characterize IVIAs in terms of their specific functions and design features that differentiate them from at-home agents. In addition, considering the dynamic function allocation between an agent and a driver as the level of automation increases, the workshop also collected a cluster of functions carried by IVIAs under the three representative levels of automation conditions: Level 0 with no driving automation, Level 3 with conditional driving automation, and Level 5 with full driving automation. In total, 30 participants from academia and industry attended the workshop from eight countries.

The present paper reports the outcomes of the workshop, focusing on discussion results and characteristics tailored for IVIAs under each of the three levels of automation conditions.

METHODS AND ACTIVITIES

As participatory design methods, expert focus groups and design activities were adopted in our workshop. Participatory design allows users and stakeholders to actively engage in the design process, which ensures that designers are addressing user problems and grounding design decisions from target stakeholders (Sharp et al., 2019). While it was challenging to involve the direct end-users in the design process for future technologies that do not exist yet, the experts and practitioners in the field of automated vehicles (AVs) could serve as subject matter experts. Thus, their opinions were valuable for us to address the challenges associated with designing IVIAs situated in future AVs.

Focus Group

On the first day of the workshop, we divided our attendees into five groups and hosted focus group discussions. The groups exchanged their ideas on four topic questions in sequence.

What makes an agent an agent? The discussion on the first question aimed to establish a definition of agents. The factors that define an agent are referred to as agent **attributes** in the following sections.

What are the special characteristics of IVIAs? After defining agent attributes, we discussed both the unique functions and design features of the agents that differentiate them as at-home agents or IVIAs. **Functions** refer to agent purpose, in other words, what the agent is used for. For instance, functions of IVIAs include but are not limited to monitoring the traffic conditions and informing drivers properly. **Design features** specify both the visual and auditory specifications of agents, such as agent appearance, agent voice, and agent facial expressions, if any. In addition to these objective features, perceived features such as agent attitudes and user trust are discussed. Functions and design features together characterize the agent into different categories: at-home agents or IVIAs.

How do drivers and passengers perceive or distinguish these two categories of agents? With specified characteristics for IVIAs, the discussion further extended to how users distinguish IVIAs from at-home agents depending on variation in functions and design features.

What are users' preferences toward the form of agents? Finally, we discussed whether users prefer a one-fits-all agent that serves users both at home and in the vehicle, or multiple agents specialized in different situations.

Design Activity

On the second day of the workshop, we had some returning attendees and new attendees different from the those in the first day of activity. They were divided into three groups to brainstorm characteristics for an agent to serve in each of the three levels of automation conditions (Level 0, Level 3, and Level 5). The group first determined agent functions and then considered design features.

DISCUSSION OUTCOMES

Affinity diagrams, the simplest way to reveal common themes across all idea contributors (Holtzblatt & Beyer, 2017), were used to organize attendees' discussion points brought out in the first-day topic discussion session.

Defining Agents

We first discussed possible definitions of agents. Both agent attributes and potential characteristics were discussed during the first-day workshop (Figure 1).

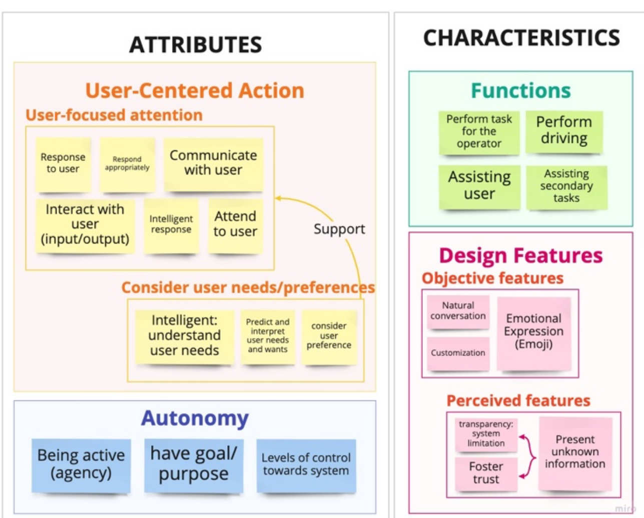

Figure 1. Affinity diagram on defining agents.

User-centered action and agent autonomy are two key attributes that qualify an interactive entity as an agent. An agent performing user-centered action will always attend to the user and pay attention to the user's status. Even though the agent cannot provide an instant response to a request, feedback on the system operating status should be provided. In addition to always attending to users, an agent can understand user needs and requirements and can consider user preferences when analyzing and responding to user requests. In fact, user-centered action is an attribute shared by all human-machine interfaces. In terms of autonomy, an agent should retain a certain level, if not all, of control over the vehicle automation system in terms of decision-making to prioritize tasks based on its objectives.

In addition to these determinative attributes, agents' functions and design features are more situated in the context than defining an agent per se. An agent can perform the primary task for users or assist with secondary tasks based on human-machine function allocation. Design features, such as communication styles and capability of customization, are also context- and task-dependent. User perception affected by the agent outputs is also an important design feature to assess agent quality. A well-designed agent can present information unknown or unclear to users, such as system limitations or external information hidden from the user, improving system transparency and fostering user trust (Chen et al., 2017; Wright et al., 2017).

Characteristics of IVIAs

Functions and design features are able to characterize agents situated in different contexts (Figure 2). IVIAs can assist users with their primary tasks, such as driving, or support them with their secondary tasks (e.g., interacting with the infotainment system) to ensure the primary task performance. In addition to functions directly related to driving support, we also discussed IVIAs' potential to monitor and intervene in driver states such as their emotions (Jeon, 2015; Jeon et al., 2014; Jeon & Walker, 2011), drowsiness (Ghizlene et al., 2019), and boredom (Samrose et al., 2020) that might influence driving performance. In fact, this type of proactive behavior essentially distinguishes IVIAs from at-home agents. Even though the IVIAs are not activated by the driver, they are proactively performing other tasks–monitoring vehicle states and evaluating road conditions–behind the scenes to guard driving safety. On the contrary, state-of-the-arts at-home agents are always passively pending users' requests.

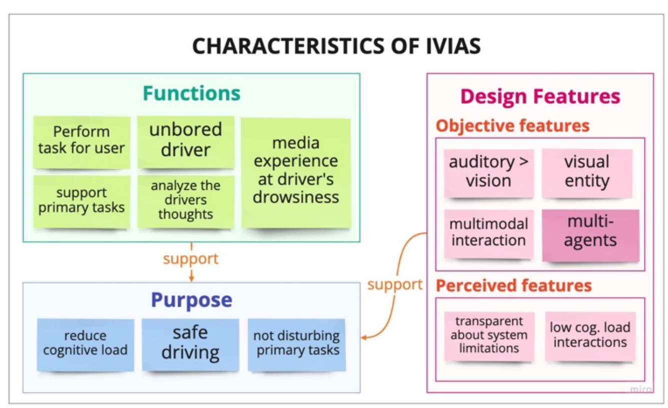

Figure 2. Affinity diagram on characteristics of IVIAs

Design features were also discussed heavily as special characteristics of IVIAs. Multimodal interaction was recommended to support driving as a visually demanding task, especially utilizing the benefits of the auditory modality to convey messages. The necessity of a visual entity for the presence of IVIAs was also discussed. It was argued that a visual presentation could be good for reference purposes, but it can also be distracting. Simplified but effective interaction is also advocated to further reduce the cognitive demand required for communicating with IVIAs. The discussion outcomes align with the continuous research effort in IVIAs. Previous research has investigated different design features to support either driving tasks or secondary tasks, or both to reduce drivers' cognitive demand and improve driving performance. For instance, various speech styles have been researched to improve takeover time (e.g., assertive vs. non-assertive (Wong et al., 2019)), reduce distraction when interacting with in-vehicle infotainment systems (Gaffar & Kouchak, 2018), or optimize driving experience (Wang, Lee, et al., 2021). Other than characteristics that support user tasks, features that facilitate user experience were also discussed as a potential for IVIAs but with less power of uniquely characterizing IVIAs. An example is the ability to customize IVIAs for multiple users. The system can take driver personalities into account to maximize user acceptance (Braun et al., 2019).

The group discussion also brought up a new research topic to be examined in the field of IVIAs. As IAs have become popular in our lives and integrated well into people's personal devices, multiple agents co-existing in the vehicle will be more common. The current knowledge available is not able to solve the problem concerning streamlining the interaction logic and resolving different agents' conflicting responses. A concept paper has also pointed out this futuristic dilemma and embodied potential challenging situations (Lee et al., 2021).

Genie vs. Jarvis

The discussion on perceiving and distinguishing at-home agents (e.g., Genie) and IVIAs (e.g., Jarvis) fell on functions and design features again (Figure 3). At-home agents and IVIAs perform different tasks in essence, with IVIAs more active in driving scenarios. Thus, users maintain corresponding expectations toward different agents. Agent voice is the most representative design feature for an agent and can differentiate one from another, aligning with the theory of computers as social actors (Nass et al., 1994).

The discussion came after distinguishing "Genie" and "Jarvis" was the issue associated with having multiple agents. Both information pieces input to and output from agents were discussed. Depending on the contexts, the agents can be initiated by voice commands at home with lower noise, or button initiation in the vehicle where voice commands can be masked, or multiple-user scenarios can confuse the agent. Agent hierarchy is also situated in the context. The agent currently performing the task should be given the priority to interpret user needs and provide the appropriate response. If none of the agents is active when requested, an initiation action is required either through physical contact (e.g., button pressed) or a specific point of reference (e.g., "Hey, Siri"). Agent-to-agent communication was also brought up and discussed in the next section.

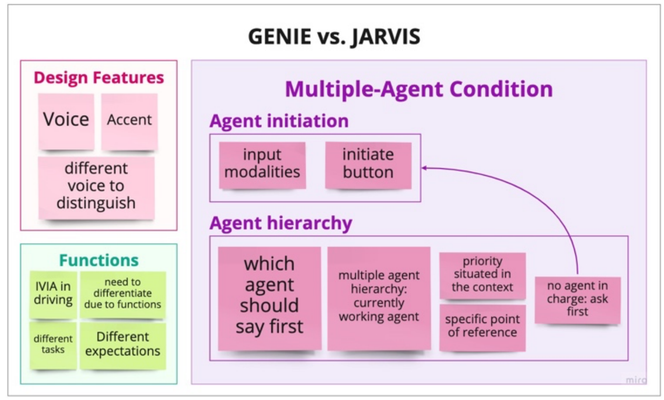

Figure 3. Affinity diagram on how to distinguish agents

One Fits All vs. Specialized Multiples

Finally, the discussion on the preference towards the one-fits-all agent or multiple specialized agents concluded on the first day of the workshop. While it highly depends on the user context, use domain, and personal taste, most attendees favored specialized multiples and debated the advantages and disadvantages of both options (Figure 4).

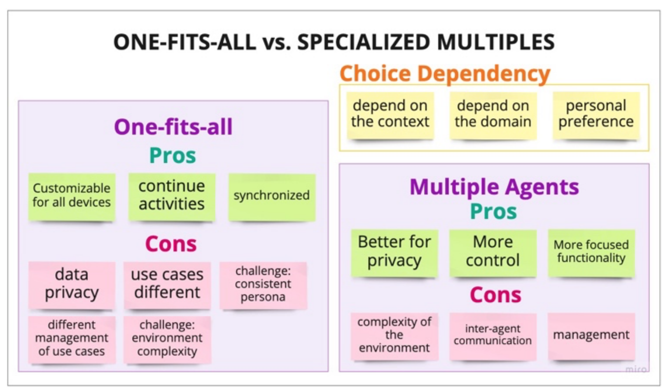

Figure 4. Affinity diagram on preferences towards agent type

Although there are challenges associated with agent management and inter-agent communication, multiple specialized agents are more focused on delivering well-designed functions, tuning sophisticated control, and providing better privacy. On the contrary, a one-fits-all design, thanks to the consistent persona established in a single system, allows synchronized information among different tasks and provides a streamlined experience in continuing the same activity across different user contexts. However, data privacy and agent complexity will become more concerned. Further investigation into this trade-off is needed to optimize futuristic user experiences in the car.

DESIGN CONSIDERATIONS FOR IVIAS

Even though IVIAs are used in the same context (i.e., driving), they can play different roles in response to different automation levels. Thus, we selected three key levels in the spectrum of automation levels (Society of Automotive Engineers, 2018) as representative use cases to fit IVIAs: Level 0 with no driving automation, Level 3 with conditional driving automation, and Level 5 with full driving automation. Three groups discussed functions and design features of each IVIA fitting for each use case. Design features considered in the design activity section of the workshop included agent form factor (e.g., physical or virtual), agent appearance, speech characteristics, speech style, conversation methods, agent attitude, affective expression, body gesture, and other features proposed by attendees. The following sections summarize the design considerations corresponding to each use case, followed by general recommendations on agent design.

IVIAs in Level 0 Automation

IVIAs in Level 0 automation are designed to reduce drivers' cognitive demand in driving-related tasks and secondary tasks. Ideally, the agent is aware of the driver's state (e.g., tiredness, anger) and can adapt accordingly.

Regarding design features, the voice-only form factor was advocated to situate in the vehicles with Level 0 automation to reduce visual distraction and avoid visual attention competition. Thus, the feedback on agent attention and operation (e.g., auditory cue) has to be properly designed to inform users about the system status. Speech characteristics, speech style, and agent attitude can be customizable and are use-context dependent, while a one-fits-all default option should also be provided.

IVIAs in Level 3 Automation

In conditionally automated vehicles, as drivers are required to take over the control when the system requests, IVIAs should timely inform drivers about upcoming events and whether the system is able to handle the event. It is also important to provide explanations regarding current vehicle behaviors and their understanding of the current situations to establish trust and promote driver-agent interaction.

The design features fitted for IVIAs in Level 3 automation are slightly different from those designed for Level 0 automation. IVIAs in semi-automated driving can have a physical body seated on the dashboard. Other features are shared with agents in Level 0, especially the safety-driving-related design considerations such as driver state monitoring and intervening. The discussion on IVIAs in Level 3 automation also brought up an important point about keeping the agent consistent but also adaptive. Essentially, features that determine the vocal timbre—such as age and gender—should be constant once set up to ensure a consistent voice-identity perception and to distinguish IVIAs from other agents (Mathias & von Kriegstein, 2019), while the speech style and speech tone can vary depending on situations to convey additional information such as urgency (Lavan et al., 2019).

IVIAs in Level 5 Automation

IVIAs in fully autonomous driving are expected to be versatile to match the advanced technology applied to automation. IVIAs are expected to cover all aspects of the in-vehicle user experience. Thus, the design features are more user-oriented to provide personalized experiences.

Attendees also discussed another design feature that helps users perceive the agent as the same or refer to the agent properly in multiple-agent conditions. Besides agent voice, agent appearance can also identify and distinguish agents.

Designing IVIAs Across Use Cases

While IVIAs serving under different levels of automaton conditions carry different objectives and functions, design considerations remained similar with slight changes to fit the functions better.

We found that agent voice and appearance are features that identify an agent or differentiate agents. Although these features should be constant once determined, other features such as speech style, tone and pitch, can vary by the situation to trigger additional perceptions such as urgency.

Based on our discussion session on the first day of the workshop, the design activity on the second day extended the key components that define and characterize the agents. While retaining the user-centered action and autonomy, IVIAs can be different in design features in various aspects. But all the variations should be carefully selected to support the functions situated under a well-scoped user scenario.

LIMITATIONS

Despite the exciting findings and promising implementations, we acknowledge that this workshop has some limitations. First, although we covered three key levels of automation in the vehicle automation spectrum, not all critical use cases were mentioned or could be discussed in the workshop. As automation advances, users can perform various tasks and their requirements can change accordingly even within the same use context. More focused workshops or discussions can help with narrowing down the design considerations for specific key use contexts. For example, working, entertaining, and relaxing can be prevalent in Level 5 Automation. Designing agents supporting these use contexts can be standalone discussions. Second, the discussions in the workshop remain at a high level. Specific design prototypes can be carried out to further ground the agent designs. However, actually prototyping the agents can be challenging through an online workshop as a result of the global COVID-19 pandemic. Thus, in-person workshops can be helpful in the future to provide a hands-on experience in designing agents.

CONCLUSIONS AND FUTURE WORKS

We presented an overview of designing intelligent agents in driving conditions. The design considerations and recommendations posed in our workshop should also be tested in experiments to prove the concepts and tune the parameters in a sophisticated approach. As our research proceeds and gains more expertise in delivering intelligent agent experience, we will continue to host the workshop to push the conceptualization and validation of IVIA design.

REFERENCE

Braun, M., Mainz, A., Chadowitz, R., Pfleging, B., & Alt, F. (2019). At Your Service: Designing Voice Assistant Personalities to Improve Automotive User Interfaces. *Conference on Human Factors in Computing Systems - Proceedings*, 1–11.

Chen, J. Y. C., Barnes, M. J., Wright, J. L., Stowers, K., & Lakhmani, S. G. (2017). Situation awareness-based agent transparency for human-autonomy teaming effectiveness. *Micro- and Nanotechnology Sensors, Systems, and Applications IX*, *10194*(May 2017), 101941V. https://doi.org/10.1117/12.2263194

Gaffar, A., & Kouchak, S. M. (2018). *Using simplified grammar for voice commands to decrease driver distraction. March*.

Ghizlene, B., Zoulikha, M., & Pomares, H. (2019). An Efficient Framework to Detect and Avoid Driver Sleepiness Based on YOLO with Haar Cascades and an Intelligent Agent. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics): Vol. 11507 LNCS*. Springer International Publishing. https://doi.org/10.1007/978-3-030-20518-8_58

Holtzblatt, K., & Beyer, H. (2017). The affinity diagram. In K. Holtzblatt & H. Beyer (Eds.), *Contextual design* (2nd ed., pp. 127–146). Elsevier. https://doi.org/10.1016/b978-0-12-800894-2.00006-5

Jeon, M. (2015). Towards affect-integrated driving behaviour research. *Theoretical Issues in Ergonomics Science*, *16*(6), 553–585. https://doi.org/10.1080/1463922X.2015.1067934

Jeon, M., & Walker, B. N. (2011). What to detect? Analyzing factor structures of affect in driving contexts for an emotion detection and regulation system. *Proceedings of the Human Factors and Ergonomics Society*, 1889–1893. https://doi.org/10.1177/1071181311551393

Jeon, M., Walker, B. N., & Yim, J. Bin. (2014). Effects of specific emotions on subjective judgment, driving performance, and perceived workload. *Transportation Research Part F: Traffic Psychology and Behaviour*, *24*, 197–209. https://doi.org/10.1016/j.trf.2014.04.003

Lavan, N., Burton, A. M., Scott, S. K., & McGettigan, C. (2019). Flexible voices: Identity perception from variable vocal signals. *Psychonomic Bulletin and Review*, *26*(1), 90–102. https://doi.org/10.3758/s13423-018-1497-7

Lee, S. C., Jeong, S., Wang, M., Hock, P., Baumann, M., & Jeon, M. (2021). To Go or Not to Go? That is the Question When In-Vehicle Agents Argue with Each Other. *Adjunct Proceedings - 13th International ACM Conference on Automotive User Interfaces and Interactive Vehicular Applications, AutomotiveUI 2021*, 223–224. https://doi.org/10.1145/3473682.3481876

March, S., Hevner, A., & Ram, S. (2000). Research Commentary: An Agenda for Information Technology Research in Heterogeneous and Distributed Environments. *Information Systems Research*, *11*(4), 327–341. https://doi.org/10.1287/isre.11.4.327.11873

Mathias, S. R., & von Kriegstein, K. (2019). *Voice Processing and Voice-Identity Recognition*. https://doi.org/10.1007/978-3-030-14832-4_7

Nass, C., Steuer, J., & Tauber, E. R. (1994). Computers are social actors. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems Celebrating Interdependence - CHI '94*, 72–78. https://doi.org/10.1145/191666.191703

Samrose, S., Anbarasu, K., Joshi, A., & Mishra, T. (2020). Mitigating Boredom Using An Empathetic Conversational Agent. *Proceedings of the 20th ACM International Conference on Intelligent Virtual Agents, IVA 2020*. https://doi.org/10.1145/3383652.3423905

Sharp, H., Preece, J., & Rogers, Y. (2019). *Interaction design: Beyond human-computer interaction* (5th ed.). John Wiley & Sons.

Society of Automotive Engineers. (2018). *Taxonomy and definitions for terms related to on-road motor vehicle automated driving systems*.

Wang, M., Hock, P., Lee, S. C., Baumann, M., & Jeon, M. (2021). Genie vs. Jarvis: Characteristics and Design Considerations of In-Vehicle Intelligent Agents. *13th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, 197–199. https://doi.org/10.1145/3473682.3479720

Wang, M., Lee, S. C., Kamalesh Sanghavi, H., Eskew, M., Zhou, B., & Jeon, M. (2021). In-Vehicle Intelligent Agents in Fully Autonomous Driving: The Effects of Speech Style and Embodiment Together and Separately. *13th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, 247–254. https://doi.org/10.1145/3409118.3475142

Wong, P. N. Y., Brumby, D. P., Babu, H. V. R., & Kobayashi, K. (2019). "Watch out!" Semi-autonomous vehicles using assertive voices to grab distracted drivers' attention. *Conference on Human Factors in Computing Systems - Proceedings*, 5–10. https://doi.org/10.1145/3290607.3312838

Wright, J. L., Chen, J. Y., Barnes, M. J., & Hancock, P. A. (2017). *Agent Reasoning Transparency : The Influence of Information Level on Agent Reasoning Transparency : The Influence of Information Level on Automation-Induced Complacency*. *June*.