The Properties of Intelligent Human-Machine Interface

Alexander ALFIMTSEV, Vladimir DEVYATKOV
Information systems and telecommunications, Bauman Moscow State Technical University
Moscow, 105005, Russia

ABSTRACT

Intelligent human-machine interfaces based on multimodal interaction are developed separately in different application areas. No unified opinion exists about the issue of what properties should these interfaces have to provide an intuitive and natural interaction. Having carried out an analytical survey of the papers that deal with intelligent interfaces a set of properties are presented, which are necessary for intelligent interface between an information system and a human: absolute response, justification, training, personification, adaptiveness, collectivity, security, hidden persistence, portability, filtering.

Keywords: Intelligent interface, human-computer interaction, artificial intelligence, modality, predicate.

1. INTRODUCTION

Today there are a lot of various definitions for intelligent human-machine interface [4]. In this paper the intelligent human-machine interface (IHMI) is assumed to be an interface that provides an interaction of user and an information system (IS) in a natural way by means of intrinsic modalities of a human (gesture, voice, facial expression) using artificial intelligence and pattern recognition methods. Intelligent human-machine interfaces based on multimodal interaction are developed separately in different application areas. No unified opinion exists about the issue of what properties should these interfaces have to provide a barrier-free interaction. Thus the purpose of this research is to formulate the major necessary properties of such interfaces.

The paper attempts to substantiate these properties using a formal model of IHMI, i. e. a mediator that has access to information about user and IS. In other words IHMI is an intellectual system-mediator based on the knowledge about behavior both the user and the IS. This work doesn't deal with particular knowledge description languages. Only some features of IHMI architecture as a knowledge-based system are considered.

This paper is organized as follows. Section 2 discusses a model of a barrier-free IHMI. In section 3 the major properties of such IHMI are described and substantiated: absolute response, justification, training, personification, adaptiveness, collectivity, security, hidden persistence, portability, filtering. Section 4 concludes the paper.

2. IHMI MODEL

IHMI is a mediator between the user and the IS. It operates on the basis of knowledge about behavior both of them. Using this knowledge IHMI has to choose a particular strategy of a barrier-free interface in a situation-dependent way. Thus at our opinion the most suitable architecture for IHMI is a multiagent architecture. In order to simplify formulation of the properties, the architecture is assumed to be reactive [2]. Finite-state automaton will be a model for each agent.

Thus a set of automata $M_1, M_2, \ldots, M_m$ operating in parallel and interacting will be a model for IHMI. Let $\boldsymbol{b}(t)=\{b_1(t),..., b_m(t)\}$ be a set of internal states of the corresponding automata $\{M_1, M_2, \ldots, M_m\}$ respectively at time instance t called an internal macro state. Let $y_j(t) = \varphi(b_j(t))$ be an output function and $b_j(t+1)=f_j(x_j(t), b_j(t))$ be a state-transition function of state machine M_j where $b_j(t)$, $b_j(t+1)$ are internal states of M_j. Let $\boldsymbol{internal}(\boldsymbol{b}(t))$ be a predicate that becomes true when the automata are in macro state $\boldsymbol{b}(t)$, $\boldsymbol{input}(\boldsymbol{x}(t))$ be a predicate that becomes true when the automata are fed with an input macro state $\boldsymbol{x}(t)$, be a predicate that becomes true when the automata yield an output macro state . Let $\boldsymbol{y}(t)=\boldsymbol{\varphi}(\boldsymbol{b}(t))$ be a output macro function, $\boldsymbol{b}(t+1)=\boldsymbol{f}(\boldsymbol{x}(t),\boldsymbol{b}(t))$ be state-transition macro function. Using this notation we will formulate the properties of IHMI in terms of modal temporal logic. We will use modal operators listed in a table below.

Operator notation	Operator meaning
○ξ	ξ has to become true next time instance
◉ξ	ξ was true last time instance
◊ξ	ξ eventually has to become true in the future
□ξ	ξ has to be true always in the future
♦ξ	ξ was true at some of the previous time instances
■ξ	ξ was always true in the past
σuξ	ξ has to be true while σ is true
σsξ	ξ has to become true as soon as σ will become true
σwξ	ξ is true when σ is true

According to the modal temporal logic philosophy one must not use time symbols explicitly in the statements. And according to Gabbay theorem [5] any statement of modal temporal logic can be transformed into a logically equivalent form using implicative rules like *past* $\supset$ *future*.

Environment the agent IHMI faces consists of users. Agents communicate with users by means of two types of predicates: predicates being sent and predicates being received. Predicates being sent are located in the antecedent part of the rule and predicates being received are located in the consequent part of the rule.

3. IHMI PROPERTIES

3.1. Absolute response property

User calls IHMI using various modalities. For the simplicity purposes from here till the end of the paper this request is assumed to occur in discrete time instances. Regardless the number of modalities used at any time instance t their values are transformed into a sequence of input macro states $\boldsymbol{x}(t)$. Such input macro state can be considered as a word of some language. Sequence of such words corresponds to a sequence in this language. If text language is used for call (e. g. DISL, GIML, ISML [6]) input state can correspond to a distinct word of this language or as well as its sentence. No matter if sensors [18], speech, gesture or graphics [11] is used for call to IHMI, a sequence of input states is assumed to be used anyway. In answer to the sequence of input macro states IHMI yields a response. This response can be expressed in the same or another language using the same or another modalities. Being an answer to the user's request, it is a sequence of output macro states.

Thus in terms of modal temporal logic absolute response property can be expressed as:

$$(\forall \boldsymbol{x})\ (\boldsymbol{input}(\boldsymbol{x}) \supset \boldsymbol{internal}(\boldsymbol{b}_i) \wedge \bigcirc(\boldsymbol{internal}(\boldsymbol{b}_j, \boldsymbol{b}_j{=}\boldsymbol{f}(\boldsymbol{x}, \boldsymbol{b}_i)) \wedge \boldsymbol{output}\ (\boldsymbol{y}_j, \boldsymbol{y}_j = \boldsymbol{\varphi}(\boldsymbol{b}_j)).$$

This property is to be read as "When an input macro state $\boldsymbol{x}$ is fed to IHMI next time instance IHMI, being in an internal macro state $\boldsymbol{b}_i$, will transit to state $\boldsymbol{b}_j$ and yield an output state $\boldsymbol{y}_j$". This statement is true for all $\boldsymbol{x}$ values, so one can test IHMI's response to any input macro state.

3.2. Justification property

To understand the way IHMI solves the given task user has the right to ask for a justification for the solution IHMI has found. For the chosen architecture such justification may be rather simple. The solution of the task is sequence of responses (output macro states) $\boldsymbol{y} = \boldsymbol{\varphi}(\boldsymbol{b})$, each one caused by a user request, i. e. a single or a temporal sequence of input macro states. So in the simplest case justification may be a temporal sequence of triples "output macro state / input macro state / internal macro state", where the output macro state is returned by the output macro function evaluated for an internal state transition caused by the given input macro state. Though the justification can be further detailed, anyway it can be represented as hierarchy of such triples "output value / input value / internal value". Such justification property is similar to explanation function, which is typical for modern expert systems [19].

Thus the justification property can be defined in terms of modal temporal logic at macro definition level as follows:

$$(\forall \boldsymbol{y})\ (\boldsymbol{ouput}(\boldsymbol{y}) \supset\ (\bullet(\boldsymbol{internal}(\boldsymbol{b}_i) \wedge \boldsymbol{input}\ (\boldsymbol{x})) \wedge \boldsymbol{internal}(\boldsymbol{b}_j, \boldsymbol{b}_j{=}\boldsymbol{f}\ (\boldsymbol{x}, \boldsymbol{b}_i)\ ) \wedge \boldsymbol{ouput}\ (\boldsymbol{y}, \boldsymbol{y} = \boldsymbol{\varphi}(\boldsymbol{b}_j).$$

This property is to be read as "If IHMI yielded an output macro state $\boldsymbol{y}$, then at previous time instance it had been in an internal macro state $\boldsymbol{b}_i$ and after an input macro state $\boldsymbol{x}$ had been fed to IHMI, it transmitted to an internal macro state $\boldsymbol{b}_j$, evaluated by the state-transition macro function $\boldsymbol{f}$ which results in an output state $\boldsymbol{y}_j$ evaluated by the output macro function $\boldsymbol{\varphi}$ of internal state $\boldsymbol{b}_j$". This statement is true for all $\boldsymbol{y}$ values, so one can find justification for any IHMI's response represented by an output macro state.

3.3. Training property

When user starts to use IS one of the major tasks for IHMI is to introduce him to the capabilities that IHMI provides. One of the popular introduction techniques is to demonstrate IS behaviors by means of IHMI. Each behavior corresponds to an automaton. One can consider each internal state of some automaton as an atomic step of behavior execution, output state as a demonstration for user of what this step results in, and input state as a user's reaction to this demonstration.

Given such state interpretation and such automaton application for behavior demonstration purposes, that enables user to respond to every atomic step, provided that all atomic steps will be demonstrated and all user responses will be performed, the training property can be defined in terms of modal temporal logic at macro definition level as follows:

$$(\forall \boldsymbol{y}_i)\ (\boldsymbol{internal}(\boldsymbol{b}_i) \wedge \boldsymbol{output}(\boldsymbol{y}_i, \boldsymbol{y}_i = \boldsymbol{\varphi}(\boldsymbol{b}_i))\ \supset \bigcirc(\ \boldsymbol{input}\ (\boldsymbol{x}) \supset \boldsymbol{internal}(\boldsymbol{b}_j, (\boldsymbol{b}_j = \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{b}_i)) \wedge \boldsymbol{output}(\boldsymbol{y}_j, \boldsymbol{y}_j = \boldsymbol{\varphi}(\boldsymbol{b}_j)).$$

Given the described state interpretation this property is to be read as "Let $\boldsymbol{b}_i$ be an atomic step of the behavior execution being demonstrated, which corresponds to an output state $\boldsymbol{y}_i$. At next time instant provided that an input macro state $\boldsymbol{x}$ is fed to IHMI as a user's response, IHMI proceeds to the atomic step $\boldsymbol{b}_j$ evaluated by state-transition macro function $\boldsymbol{f}$, and a new output state is evaluated as a function $\boldsymbol{\varphi}$ of state $\boldsymbol{b}_j$". It should be noted that idle is also considered as user reaction.

3.4. Personification property

This property refers to an interface adjustment according to the user preferences. For the chosen architecture it means that automaton behavior should be adjusted to best meet user behavior features during interaction with IHMI. For instance, some automaton transitions, which have been inactive for a long time, may be disabled or completely removed. One can consider the personification property of IHMI as its ability to learn how to interact with a particular user [3, 17].

For instance, the personification property can be defined in terms of modal temporal logic at macro definition level as follows:

$$(\forall \boldsymbol{x}, \boldsymbol{b}_i)\ \blacklozenge(\boldsymbol{input}(\boldsymbol{x}) \wedge \boldsymbol{internal}(\boldsymbol{b}_i)) \supset \boldsymbol{input}(\boldsymbol{b}_j, \boldsymbol{b}_j = \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{b}_i))\, u\ \boldsymbol{output}(\boldsymbol{y}_j, \boldsymbol{y}_j = \boldsymbol{\varphi}(\boldsymbol{b}_j)).$$

The meaning of this property consists in the following. As far as new input macro states $\boldsymbol{x}$ are provided when IHMI is in the macro state $\boldsymbol{b}_i$ and transition to the macro state $\boldsymbol{b}_j$ is valid, the output macro state $\boldsymbol{y}_j$ is also valid. Otherwise it is invalid. Using modal operators one can derive other personification properties which in fact are the restrictions for behavior of IHMI automaton model.

3.5. Adaptiveness property

Unlike the personification property, that permanently constrains behavior of IHMI model, the adaptiveness property permits reconfiguration of the IHMI model by means of adding or deleting some states and transitions depending on the user's behavior and other factors. Commonly it requires input data analysis [1, 20] or user behavior monitoring.

For instance, the following property denotes, that since IHMI fed with input state $\boldsymbol{x}$ does not enter the internal state $\boldsymbol{b_i}$ any more, transition $\boldsymbol{b_j} = \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{b_i})$ and output state $\boldsymbol{y}_j = \boldsymbol{\varphi}(\boldsymbol{b_j})$ have become invalid.

$$(\forall \boldsymbol{x}, \boldsymbol{b}_i) \blacklozenge \neg(\boldsymbol{input}(\boldsymbol{x})) \wedge \boldsymbol{internal}(\boldsymbol{b}_i) \wedge s\, \neg(\boldsymbol{internal}(\boldsymbol{b_j}, \boldsymbol{b_j} = \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{b}_i)) \wedge \boldsymbol{output}(y_j, \boldsymbol{y}_j = \boldsymbol{\varphi}(\boldsymbol{b_j})).$$

3.6. Collectivity property

IHMI collectivity means its ability to provide access to the same data, e. g. visual data [9], for several users simultaneously. Application of such interface is especially important in operations room, e. g. for decision making about elimination of the consequences of the accident, when cross-function team is involved. In this case IHMI has to provide a coordinated decision making by a group of users [21]. Many corresponding strategies are considered in game theory and artificial intelligence. Within the scope of this paper it is impossible to discuss and define all of them in terms of modal logic.

The simplest strategy, called Nash equilibrium [13], is considered below as an example. According to this strategy, in case of two users their behavior can be described by the following rules:

- when user i makes decision $\boldsymbol{x}_i$, user j can do nothing better than to make a decision $\boldsymbol{x}_j$;
- when user j makes decision $\boldsymbol{x}_j$, user i can do nothing better than to make a decision $\boldsymbol{x}_i$;

In terms of modal logic this strategy as a property can be written in the following way:

$$[(\forall \boldsymbol{x_i}, \boldsymbol{b})\ (\boldsymbol{input}(\boldsymbol{x}_i)) \wedge \boldsymbol{internal}(\boldsymbol{b}) \supset \boldsymbol{input}(\boldsymbol{x_j})] \wedge$$
$$[(\forall \boldsymbol{x}_j, \boldsymbol{b})\ (\boldsymbol{input}(\boldsymbol{x_j})) \wedge \boldsymbol{internal}(\boldsymbol{b}) \supset \boldsymbol{input}(\boldsymbol{x_i})].$$

The rest IHMI properties this paper considers require detailed IHMI model description to define them in terms of modal logic, which is impossible in one paper. Therefore only a common description is regarded.

3.7. Data security property

Data security is a highly investigated problem, which assumes that user may forget to save the data, accidentally delete it or send it to a wrong address, etc. Thus IHMI has to inform user about all his mistaken actions during data processing. Description of security properties in terms of modal logic depends on particular data security requirements.

3.8. Access security property

Access security is a highly investigated problem too. IHMI security features like biometric authentication are able to provide high access security level. Their description in terms of modal logic requires even more detailed definition of all relevant processes.

3.9. User security property

User security from the IHMI point of view is closely associated with a particular IS application domain. For instance, in medical IS sometimes it is quite enough to report blood pressure and heart rate to user in order to prevent unwanted sequelae and to transmit this data to the emergency department.

3.10. Hidden persistence property

This property denotes the interface ability to focus user's attention upon first-priority tasks: walk, conference conduction or any other activity that requires full attention. Under such conditions IHMI has to non-importunately guide user's behavior. For instance, while user is walking on the stairs, IHMI can display a building plan on the walls around him. Nowadays such technologies become more and more widespread [12]. At present hidden persistence properties of IHMI also appear in the form of operational and interoperational persistence [4]. Operational persistence means that interface is always ready to interact with user. Interoperational persistence means that the main functionality of interface is always accessible for user.

3.11. Portability property

Regardless of the application domain intelligent interface has to be portable. This assumes its ability to interact with other IS, using modern data communication protocols [8, 7].

3.12. Filtering property

On the one hand, this property [16] makes it possible to reduce the amount of data the user analyzes while seeking for a problem solution. On the other hand, it enables IHMI to find and report an exclusive information that user couldn't find himself because of his incapacity to view great amounts of data and to detect complicated regularities in data.

4. CONCLUSION

A primary goal of intelligent interfaces is to support natural, efficient, powerful, and flexible human-computer interaction. If the interaction technology is awkward, or constraining, the user's experience with the synthetic environment is severely degraded. If the interaction itself draws attention to the technology, rather than the task at hand, it becomes an obstacle to a successful virtual environment experience.

The traditional two-dimensional, keyboard- and mouse-oriented graphical user interface is not well-suited for virtual environments. Instead, synthetic environments provide the opportunity to utilize several different modalities and integrate them into the user experience. Therefore in this paper the main properties of intelligent human-machine interface are considered. Its main advance consists in attempt to define these properties in terms of modal temporal logic and thus start the process of their understanding and formalizing.

Acknowledgments. This work was supported in part by Grant of Russian Federation President (project № MK-3013.2010.9).

5. REFERENCES

[1] A. Alfimtsev, V. Devyatkov "Dynamic Gesture Recognition Using Fuzzy Model", **Proc. of WMSCI**, Orlando, USA, 2009.

[2] R. A. Brooks, **Cambrian Intelligence**, MIT Press, Cambridge, MA, 1999.

[3] Z. Duric, W. Gray, R. Heishmen, F. Li , A. Rosenfeld, C. Schunn, H. Wechsler, "Integrating perceptual and cognitive modeling for adaptive and intelligent human-computer interaction", **PIEEE**(90), No. 7, 2002, pp. 1272-1289.

[4] P. Ehlert, "Intelligent user interfaces: introduction and survey", **Research report** DKS03-01, the Netherlands, 2003, pp. 20-49.

[5] D. Gabbay, "Craig Interpolation Theorem for Intuitionistic Logic and Extensions", **Journal of Symbolic Logic**, 42 (2), 1977, pp. 269-271.

[6] J. Guerrero-Garcia, J. Munoz-Arteaga, "A Theoretical Survey of User Interface Description Languages: Preliminary Results", **LA-WEB**, 2009, pp. 36-43.

[7] R. Harper et al, "Being human: Human-computer interaction in the year 2020", **Microsoft research** Ltd, Cambridge, 2008, pp. 32-51.

[8] F. He, A. Agah, "Multi-modal human interactions with an intelligent interface utilizing images, sounds and force feedback", **Journal of intelligent and robotics systems**, №32, 2001, pp. 171-190.

[9] I. Jung et al, "Intelligent agent based graphic user interface for e-physician", **Engineering and Technology**, №36, 2007, pp. 194-197.

[10] J. E. Lindop, G. M. Treece, A.H. Gee, R.W. Prager, "An intelligent interface for freehand strain imaging", **Ultras. in med. and Bio.**, 34 (7), 2008, pp. 1117-1128.

[11] M. T. Maybury, "Language Technology: A Survey of the State of the Art Language Resources", http://www.mitre.org/work/tech_papers, 2002, 24 p.

[12] S. Mann, **Intelligent Image Processing**, NY:Wiley-Interscience, 2002.

[13] M. J. Osborne, A. Rubinstein, **A Course in Game Theory,** Cambridge, MA: MIT, 1994.

[14] D. A. Pospelov, "Intelligent interfaces for a new generation computers", **Electronnaya vichislitelnaya technika**. 1989. №3, pp. 4-20.

[15] E. V. Popov, "Communication with computer in human language", Moscow: **Nauka**, 1982, pp. 11-33.

[16] D. Raskin, **The Humane Interface: New Directions for Designing Interactive Systems**, Addison-Wesley, 2000.

[17] A. L. Rongin, A. A. Karpov, I. V. Li, "Speech and multimodal interfaces", Moscow: **Nauka**, 2006, 172 pp.

[18] H. Shinjo, U. Yamaguchi, A. Amano, K. Uchibe, A. Ishibashi, H. Kuwamoto, "Intelligent user interface based on multimodal dialog control for audio-visual systems", **Hitachi Hyoron**, v. 87, №10, 2005, pp. 801-804.

[19] V. L. Stephnjuk, A. V. Jojikashvili, "Cooperative computer", Moscow: **Nauka**, 2007, pp. 215-232.

[20] S. F. Wong, R. Cipolla, Continuous Gesture Recognition using a Sparse Bayesian Classifier, **Proc of 18th International Conf. on Pattern Recognition**, 2006, pp. 1084-1087.

[21] G. Ye, "Applying Vision to Intelligent Human-Computer Interaction", **Doctoral Thesis**, The Johns Hopkins University, 2005, pp. 4-72.