A GENERIC FORMAL SPECIFICATION OF FUSION OF MODALITIES IN A MULTIMODAL HCI

Yamine AIT AMEUR[1] and Nadjet KAMEL[1,2]

LISI/ENSMA-Université de Poitiers[1] *LRIA-USTHB*[2]
Site du Futuroscope. 86960 Futuroscope Cedex. *BP 32. El Allia. 16111 Algiers.*
France *Algeria*
{yamine,kamel}@ensma.fr

Abstract   This paper is an overview of a generic formal description allowing to encode multi-modal interactive systems, their behaviors and properties.

Keywords: Multi-modal Interaction, Formal Methods, Transition Systems.

1.   Introduction

The development of computers on the one hand, and of input and output devices on the other hand allow new system interaction modes. Indeed, nowadays voice, touch, movements, etc. can be used in order to interact with a given system. The use of such interaction modes increases the usability and the ease of use of a given system. Moreover, these interaction modes are close to those used by human beings. However, handling these new interaction modes has favored the emergence of a number of problems like fusion of input interaction, fission of output interaction, modelling, describing, designing and coding of multi-modal systems, etc.

The objective of this paper is to overview a formal description technique of a multi-modal interaction system in order to help the designers to describe, design, validate and check the properties of a multi-modal interaction system using particular formal methods or techniques. We propose a formal description based on transition system, independent of any type of interaction modes and of any particular multi-modal system which helps the designers in describing either a multi-modal system or the system required properties.

2.   Multi-modal HCI

Multi-modal interaction is complex since it supports complex events issued from different input channels. As a consequence, parallelism with all its diffi-

culties is induced. Several modalities (each modality has several basic events), may be used in order to build an interaction with the system. The building process associated to the interaction is based on a composition of sub-interactions which are themselves compositions of sub-interaction tasks and so on, until basic events are reached. Composition operators are needed in order to build such multi-modal interactions.

Multi-modal interactive systems can be categorized following several criteria. According to (Nigay and Coutaz, 1993, Nigay and Coutaz, 1995, Bellik, 1995), we base our classification on three criteria: (1) production of interactions either in parallel or in sequence; (2) use of medias exclusive or simultaneous, and; (3) number of medias per interaction. The combination of these criteria leads to seven types of multi-modal interactions.

1- *Exclusive.* Only one modality is used and the interaction are produced sequentially.

2- *Alternate.* Several modalities can be used alternatively and interactions are produced sequentially.

3- *Synergistic.* Interactions are produced concurrently but events of different modalities can be fired in parallel or interleaved.

4- *Parallel exclusive.* Interactions are produced concurrently and only one event of a modality is fired.

5- *Parallel simultaneous.* Several independent interactions can be produced concurrently and only one modality is used for each interaction. The events of the modalities may be fired concurrently.

6- *Parallel alternate.* Several independent interactions can be produced concurrently and several modalities may be used to produce an interaction but only one modality is active each time.

7- *Parallel synergistic.* Several independent interactions can be produced concurrently, several modalities may be used to produce an interaction and several events may be fired each time.

Formal methods in Multi-modal HCI. As outlined in (Palanque and Schyn, 2003), only few work for applying formal methods for the development of multi-modal interactive systems have been achieved. Three main approaches may be distinguished. First the approach of (Paterno and Mezzanotte, 1994) uses interactors to model the application of the *matis* case study starting from a task model described in a user task notation (UAN). The Lotos formal technique has been used to encode these interactors and the ACTL temporal logic supported by the Lite model checker tool has been used for verifying properties expressed on these interactors. The second approach is due to (Duke and Harrison, 1993, Duke and Harrison, 1995) (MacColl and Carrington, 9908). They show how formal techniques, based on proof systems, may be used to encode a multi-modal application. Finally, the third approach of

the Interactive Cooperative Objects (ICO) of (Palanque et al., 1995, Palanque and Schyn, 2003) based on Petri nets deals with fusion of modalities. These approaches have a partial coverage of the engineering of multi-modal interaction. Indeed, some of them deal with system representation, or with fusion of modalities while others address the property verification and validation.

We observe that all the previous approaches use specific formal techniques that have their own domain of efficiency. None of them supports the whole multi-modal interactive system design.

3.   A generic representation of Multi-modal HCI

Our approach does not consist in the use of a formal technique in order to design a multi-modal interactive system. It suggests a formal methodology allowing to represent a multi-modal interactive system independently of any particular formal technique. This approach is based on the expression of the system and of the properties corresponding to the user requirements. Our proposal consists in describing both the system and the properties in a generic and universal formal description technique: transition systems for the system representation and logics for the properties expression. Notice, that we do not push any particular formal technique nor a particular tool. Moreover, we do not give any recommendation about the way the system is designed. Indeed, at least three scenarios can be identified (1) describe the system first and then the properties it shall satisfy; (2) describe the properties and extract a system that fulfills these properties (3) describe in parallel the system and the properties to be satisfied. The scenario is usually imposed by development practices or methods and by the chosen development technique.

System description: formal specification.   The chosen representation for a multi-modal interactive system is based on the theory of interactive systems and of process algebra developed by several authors. Transition systems encoding the interactive system are used to represent formally this system. This formal description technique is universal and well understood. Moreover, several semantic aspects can be encoded using this description technique. One can encode synchrony/asynchrony, parallelism/interleaving and sequentiality. It is up to the designer to choose the semantics he/she thinks to be well adapted to his/her description of the problem. This possibility increases the description power of the system.

Finally, several composition/decomposition operations have been developed on transition systems based description. Indeed, refinement, synchronous products, abstraction, etc. operations have been formally defined. They allow to structure the development of a given system using a compositional approach providing ascending or descending developments.

418

Expression of properties. (Coutaz et al., 1995) have identified the properties that need to be expressed in order to assert that a multi-modal interactive system is usable. The authors have defined Complementarity, Assignment, Redundancy and Equivalence as the main properties that may be satisfied by a multi-modal system.

Properties are formal representations of intended behavior of a given system. Several representations of properties can be suggested. In our approach for the description of multi-modal interactive systems, we focus on two representations commonly available when using formal techniques. The first one consists in expressing properties, which are checked on the transition system describing the system, by logics whose semantics is given in terms of transition systems. Model checking techniques and proof based techniques are based on such an approach. The second approach for properties verification is based on behavioral descriptions of properties. Indeed, properties are described by transition systems that describe a suited behavior of the system. Language inclusion, simulation, bi-simulation relationships are used to check that the described behavior is also a possible behavior of the system. Proof based and model checking approaches may be used to establish such properties.

Methodology. Our proposal consists in a global generic model for handling multi-modal interactions. We focus on input modalities and their fusion. This approach consists in defining a formal representation for both the system and the properties to be expressed on the system. The expression of the system is based on general transition systems while properties are expressed either by transition systems expressing suited behaviors or by logical expression in a given logic. We just use basic formal description techniques and leave the choice of a formal technique to the methodology a designer wants to use.

System description. The syntax of the language describing the input multi-modal interactions is given by the following grammar issued from classical process algebra. The rule defining S generates the user task interactions at a higher level. These tasks use basic interaction events of the set $A = \bigcup_{i=1}^{n} A_{m_i}$ where A_{m_i} is the set of events e_i produced by a modality m_i. We denote by A_{Set} any subset of A.

$S ::= S[]S \mid S >> S \mid S\|\|S \mid S\|\|S \mid E$	higher order multi-modal interaction.
$E ::= e; E \mid e\|\|\|E \mid e\|\|E \mid \delta$ with $e \in A$	basic events rules

where $[], >>, \|\|\|, \|\|$ and ; stand for choice, enabling, interleaving, parallelism and sequence operators.

Formal semantics of the system. The underlying semantics of a multi-modal system with input modalities is a transition system. Let P and Q be two terms of the previous grammar and e, e_1 and e_2 be events of $A,$ then the tran-

sition $P \xrightarrow{e} Q$ expresses that the term Q is obtained from P when the event e occurs. Using, this notation for transitions, the operational semantics is formally expressed by transition rules expressing the behavior of each operator of the previously described system. According to Plotkin, 1981, each rule of the form $\frac{premises}{conclusion}$ expresses that when the premises hold, then the conclusion holds. The formal semantics is given by the following set of rules.

δ : empty axiom	$\delta \not\xrightarrow{}$	
; : prefix axiom	$e; P \xrightarrow{e} P$	
[] : rules	$\dfrac{P\xrightarrow{e}P'}{P[]Q\xrightarrow{e}P'}$	$\dfrac{Q\xrightarrow{e}Q'}{P[]Q\xrightarrow{e}Q'}$
>> : sequence rules	$\dfrac{P\xrightarrow{e}P' \ and \ P'\neq\delta}{P>>Q\xrightarrow{e}P'>>Q}$	$\dfrac{P\xrightarrow{e}P' \ and \ P'=\delta}{P>>Q\xrightarrow{e}Q}$
||| : interleaving rules	$\dfrac{P\xrightarrow{e}P'}{P\|\|\|Q\xrightarrow{e}P'\|\|\|Q}$	$\dfrac{Q\xrightarrow{e}Q'}{P\|\|\|Q\xrightarrow{e}P\|\|\|Q'}$
|| : parallel rules	$\dfrac{P\xrightarrow{e}P'}{P\|\|Q\xrightarrow{e}P'\|\|\|Q}$	$\dfrac{Q\xrightarrow{e}Q'}{P\|\|Q\xrightarrow{e}P\|\|\|Q'}$
	$\dfrac{P\xrightarrow{e_1}P',Q\xrightarrow{e_2}Q' \ with \ e_1\in A_{m_i},e_2\in A_{m_j},A_{m_i}\neq A_{m_j}}{P\|\|Q\xrightarrow{(e_1,e_2)}P'\|\|Q'}$	

These rules will be encoded according to the chosen formal used technique.

Representation of multi-modal interactions.   The previous model has shown a notation allowing to represent any multi-modal input interaction expressed by a user in order to fire a given action of the target system. However, usually, not all the composition operations are allowed to combine different modalities. This section shows that it is possible to extract different subsystems corresponding to particular multi-modal interaction systems. In the following we define the restricted interaction elements allowed for particular by restricting the allowed syntax.

Exclusive multi-modal interaction.

$S ::= S[]S \mid S >> S \mid E$	choice and enabling of sub-interactions.
$E ::= e; E \mid \delta$ with $e \in A_{m_i}$	events issued from one modality.

Alternate multi-modal interaction.

$S ::= S[]S \mid S >> S \mid E$	choice and enabling of sub-interactions.
$E ::= e; E \mid e\|\|\|E \mid \delta$ with $e \in A_{Set}$	events are issued from different modalities.

Synergistic multi-modal interaction.

$S ::= S[]S \mid S >> S \mid E$	choice and enabling of sub-interactions.
$E ::= e; E \mid e\|\|\|E \mid e\|\|E \mid \delta$ with $e \in A_{Set}$	events are issued from different modalities.

Parallel exclusive multi-modal interaction.

$S ::= S[]S \mid S >> S \mid S\|\|\|S \mid E$	choice, enabling and interleaving of sub-interactions.
$E ::= e; E \mid \delta$ with $e \in A_{m_i}$	events issued from one modality.

Parallel simultaneous multi-modal interaction.

$S ::= S[]S \mid S >> S \mid S\|\|S \mid E$	choice, enabling and parallelism of sub-interactions.
$E ::= e; E \mid \delta$ with $e \in A_{m_i}$	events issued from one modality.

Parallel alternate multi-modal interaction.

$S ::= S[]S \mid S >> S \mid S\|\|\|S \mid E$	choice, enabling and interleaving of sub-interactions.
$E ::= e; E \mid e\|\|\|E \mid \delta$ with $e \in A_{Set}$	events are issued from different modalities.

420

Parallel synergistic multi-modal interaction.

$S ::= S[]S \mid S >> S \mid S\|\|S \mid S\|\|S \mid E$	all operators allowed for sub-interactions.
$E ::= e; E \mid e\|\|E \mid e\|\|E \mid \delta$ with $\delta \in A_{Set}$	events are issued from different modalities.

These different definitions characterize behaviors or the multi-modal system. Properties may be expressed as well. Both behaviors and properties shall be checked using the chosen formal technique.

4. Conclusion

This paper presented a generic representation of multi-modal interactive systems. The developed approach is based on the expression of the multi-modal systems using transition systems and their associated semantics. The properties are represented either by other transition systems expressing behaviors or using a logic. This representation is independent from any particular formal technique or tool. The methodology suggests to encode the system and properties description in the formal technique the designers may use.

References

Bellik, Y. (1995). *Interfaces Multimodales : concepts, modèles et architecture.* PhD thesis, LIMSI- Université d'Orsay.

Coutaz, J., Nigay, L., Salber, D., Blandford, A., May, J., and Young, R. (1995). Four easy pieces for assessing the usability of multimodal interaction: the CARE properties. In *Proceedings of Human Computer Interaction - Interact'95,* pages 115–120. Chapman and Hall.

Duke, D. and Harrison, M. D. (1993). Abstract Interaction Objects. In *Proceedings of Eurographics conference and computer graphics forum,* volume 12, pages 25–36.

Duke, D. and Harrison, M. D. (1995). Event Model of Human-System Interaction. *IEEE Software Engineering Journal,* 10(1):3–10.

MacColl and Carrington, D. (19908). *Testing MATIS: a case study on specfication based testing of interactive systems.* FAHCI ISBN 0-86339-7948.

Nigay, L. and Coutaz, J. (1993). A design space for multi-modal interfaces: concurrent processing and data fusion. In *Proceedings of INTERCHI-93 - INTERCHI-93,* pages 172–178. ACM Press.

Nigay, L. and Coutaz, J. (1995). A Generic Platform for Adressing the Multimodal Challenge. In *Proceedings of CHI-95 - CHI-95,* pages 98–105. ACM Press.

Palanque, P., Bastide, R., and Sengès, V. (1995). Validating interactive system design through the verification of formal task and system models. In Bass, L. J. and Unger, C., editors, *IFIP TC2/WG2.7 Working Conference on Engineering for Human-Computer Interaction (EHCI'95),* pages 189–212, Grand Targhee Resort (Yellowstone Park), USA. Chapman & Hall.

Palanque, P. and Schyn, A. (2003). A Model-based for Engineering Multimodal Interactive Systems. In *9th IFIP TC13 International Conference on Human Computer Interaction (Interact '2003).*

Paterno, F. and Mezzanotte, M. (1994). Analysing MATIS by Interactors and ACTL. Technical report, Amodeus Esprit Basic Research Project 7040, System Modelling/WP36.

Plotkin, G. (1981). A Structural Approach to Operational Semantics. Technical report, Department of of computer Science, University of Arhus DAIMI FN 19.