Formal Modelling of Output Multi-Modal HCI in Event-B: Modalities and Media Allocation

Linda Mohand Oussaid, Yamine Aït-Ameur, Idir Ait Sadoune, Mohamed Ahmed-Nacer

▶ **To cite this version:**

Linda Mohand Oussaid, Yamine Aït-Ameur, Idir Ait Sadoune, Mohamed Ahmed-Nacer. Formal Modelling of Output Multi-Modal HCI in Event-B: Modalities and Media Allocation. AAAI Spring Symposium (AAAI 2014), Association for the Advancement of Artificial Intelligence; Stanford University Computer Science Department, Mar 2014, Palo Alto, United States. pp.38-43. hal-04103290

HAL Id: hal-04103290
https://hal.science/hal-04103290

Submitted on 23 May 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Formal Modelling of Output Multi-Modal HCI in Event-B: Modalities and Media Allocation

Linda Mohand-Oussaid[1], Idir Ait-Sadoune[2], Yamine Ait-Ameur[3] and **Mohamed Ahmed-Nacer[4]**
LIAS - ENSMA, Poitiers, France, CERIST, Algiers, Algeria, mohandl@ensma.fr[1]
SUPELEC, Gif-Sur-Yvette, France, idir.aitsadoune@supelec.fr[2]
IRIT - ENSEEIHT, Toulouse, France, yamine@enseeiht.fr[3]
LSI - USTHB, Algiers, Algeria, anacer@cerist.dz[4]

Abstract

Multi-modal Human-Computer Interfaces (HCI) allow interface designers to combine interactive modalities in order to increase interactive systems robustness and usability. In particular, output multi-modal interfaces allow the system to return the information to the user by using several modalities and media. In order to design such interfaces for critical systems, we proposed a generic formal model for the design of output multi-modal interfaces. The proposed model formally specifies an output multi-modal interface and enables properties verification. It consists of two successive models: the fission model that describes the semantic decomposition of information produced by the functional core into elementary information, and the allocation model that specifies the allocation of modality/media pairs for each elementary information. In a previous work (Mohand-Oussaid, Ait-Sadoune, and Ait-Ameur 2011), we have proposed an Event-B implementation of the fission model. In this paper, we present an Event-B formalization of the allocation model.

Introduction

The diversity of Human-Computer Interfaces (HCI) and the progress made in the definition of new interaction devices have led to complexity of the design and implementation steps of this kind of interfaces. The use of design patterns and HCI description notations becomes essential to master this complexity. In order to improve their flexibility and usability, many studies have proposed techniques, notations and methods for different HCI development steps, most existing approaches addressed the input multi-modality (Palanque and Schyn 2003; Kamel and Ait-Ameur 2007; Rousseau, Bellik, and Vernier 2005; Mohand-Oussaid et al. ; Navarre et al. 2005; Bouchet et al. 2008). When output multi-modal HCI are used in critical domains, their development requires rigorous design methods as well as for the functional core. The existing design approaches (Bordegoni et al. 1997; Rousseau, Bellik, and Vernier 2005) become insufficient due to their lack of formalization. Our work proposes a formal framework for the design of output multi-modal interactive systems. In (Mohand-Oussaid, Ait-Ameur, and Ahmed-Nacer 2009), we have proposed a formal model for the output multi-modal interfaces that is based on the WWHT (What, Which, How, Then) semi-formal design model (Rousseau, Bellik, and Vernier 2005). The proposed formal model consists of two models: the first model specifies the semantic fission step. The information produced by the functional core is decomposed into elementary information. The second model formalizes the modality/media allocation associated to the obtained elementary information. An Event-B implementation of the semantic fission model is presented in (Mohand-Oussaid, Ait-Sadoune, and Ait-Ameur 2011). This paper is devoted to the formalization of the allocation model, an Event-B model of the allocation step is presented in it.

Copyright © 2014, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

Output multi-modal human-computer interface formal model

When the designer needs to guarantee a safe design of a multi-modal HCI depending on semi-formal specifications, and to validate functional or usability properties, the existing models are insufficient and inadequate. To overcome this drawback, we have proposed a generic formal model for handling the description of the output multi-modal HCI design (Mohand-Oussaid, Ait-Ameur, and Ahmed-Nacer 2009). The proposed model is based on the WWHT model, it expresses the output multi-modal HCI within a formal framework (syntax, static and dynamic semantics). Indeed, it formally describes the construction of the output multi-modal HCI (multi-modal presentation) according to the designer's interface choices (see Fig.1). This multi-modal presentation is afterward instantiated by determining its lexico-syntactic contents (e.g. text to display) and morphological attributes (e.g. text font and size) in order to return it to the user. The multi-modal presentation is obtained by two successive decompositions of the output information generated by the functional core. Therefore, two formal models compose our global model.

1. The semantic fission model expresses the semantic fission or decomposition of the information generated by the functional core into elementary information units to be delivered to the user. These elementary information are linked by temporal and semantic operators.

2. The allocation model formalizes the multi-modal pre-

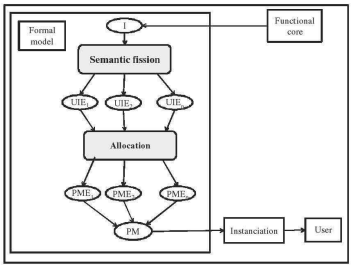

Figure 1: The output multi-modal HCI formal model

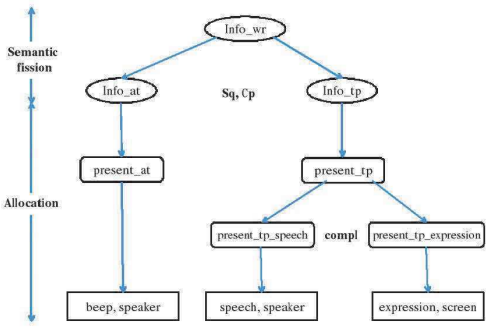

Figure 2: The temperature warning modelling process

sentation building for elementary information units resulting from the fission process. The multi-modal presentation corresponds to (modality, media) pairs, combined by the complementary, redundant, choice and iteration operators. This paper is devoted to the formal modelling of this part.

Case study. In order to illustrate our formal model, we use the *temperature warning* interface as example of output multi-modal HCI. It generates a warning message: *Abnormal temperature over 50°C*. This message is issued from the computations performed in the application functional core.

The modelling process of output multi-modal interface producing the warning message (Fig.2) involves two steps.
1. Semantic fission decomposes the information $info\_wr$ expressing *"Abnormal temperature over 50°C"* into two sequential (Sq) and complementary (Cp) elementary information units: $info\_at$ expressing *"attention"* and $info\_tp$ expressing *"temperature over 50°C"*. This means that $info\_wr$ is returned by the sequential production of the complementary information $info\_at$ and $info\_tp$.
2. Allocation builds multi-modal presentations $present\_at$ and $present\_tp$ produced from $info\_at$ and $info\_tp$. $present\_tp$ is decomposed into two complementary ($compl$) units $present\_tp\_speech$ and $present\_tp\_expression$. Finally, elementary presentation units $present\_at$, $present\_tp\_speech$ and $present\_tp\_expression$ are allocated with the corresponding pairs (modality/media) (Fig.2).

The allocation model

The allocation model is the second model in our proposed formal model. It is obtained after the fission model. For each elementary information produced by the semantic fission model, the allocation model builds the corresponding multi-modal presentation. This presentation is a combination of one or more elementary multi-modal presentation units. Each elementary multi-modal presentation unit consists of a pair (modality, media) assigned to the elementary information unit. The composition operators: complementary, redundant, choice and iteration have been identified from the WWHT semi-formal design model. They are formalized below with a syntax, a static and a dynamic semantics.

Syntax

The allocation formal model is expressed using three abstract syntactic rules. A first glue rule defines the allocation of an elementary multi-modal presentation (pme) to an elementary information unit (uie). The second one describes the decomposition of pme into elementary presentations units (upe) and the third glue rule describes the allocation of pairs (modality, media) to upe.

Let PME be the set of elementary multi-modal presentations, UIE the set of elementary information units. The first glue rule expresses the allocation to each elementary information unit, an elementary multi-modal presentation by means of the allocation function ALL. It's defined by:

$$PME ::= ALL(UIE)$$

Let UPE be the set of elementary multi-modal presentation units. The syntax of the PME expression is defined by:

$$PME ::= UPE \mid compl(UPE, PME) \mid redun(UPE, PME) \mid choice(UPE, PME) \mid iter(n, PME) \; with \; n \in N$$

Where $compl$, $redun$, $choice$ and $iter$ express respectively the complementary use of two presentations to return the elementary information uie, the redundant production of two presentations to return the elementary information uie, the alternative production of two presentations to return the elementary information uie and the n times iteration of an elementary presentation pme. Let MOD be the set of output modalities, MED the set of output media, $ITEM$ the set of $(modality, media)$ pairs such as the modality can be returned by the media, AFF a function that assigns to each UPE its corresponding $ITEM$. The allocation of modalities and media to the elementary presentation units UPE is defined by:

$$ITEM ::= AFF(UPE)$$

Static and dynamic semantics

The static semantics describes the presentation static properties during the allocation process. First, it defines the representational interpretation of presentations and their temporal boundaries. Second, it expresses properties defined on the syntactic model elements in order to describe interface robustness and usability properties. An example of such properties is media shareability property that expresses whether

the media is shareable or not. A screen is a shareable media. It may be used to produce, at the same time, more than one presentation, unlike a speaker which is a non shareable device. In order to express this property, we introduce the $share$ static property defined as function: $share : MED \rightarrow BOOL$. Thus, the shareability property is formalized by:

$$med_i \in MED \; is \; shareable \; \Rightarrow share(med_i) = true$$

The dynamic semantics addresses temporal and semantic relationships between combined presentations using static definitions. It expresses semantic combination of presentations (complementarity, redundancy) and defines temporal scheduling of presentations for the choice and iteration combination. For example, dynamic semantics for choice combination operator is defined as follows:

$$choice(p_i, p_j) = p_i \text{ if } random = 1$$
$$choice(p_i, p_j) = p_j \text{ if } random = 0$$

Application to the defined case study

We describe in the following the allocation process for the elementary information unit $info\_tp$ which expresses: "temperature over 50°C . Its corresponding allocated presentation is $present\_tp$, it's produced by means of a chatter robot dialog.

$$present\_tp = ALL(info\_tp)$$

The elementary presentation $present\_tp$ is decomposed in the allocation process into two complementary presentations units: $present\_tp\_speech$ and $present\_tp\_expression$. The first one expresses the chatter robot speech and the second one it's facial expressions : $present\_tp = compl(present\_tp\_speech, present\_tp\_expression)$

Finally, the two presentations units $present\_tp\_speech$ and $present\_tp\_expression$ are assigned with their corresponding pairs (modality, media). The $present\_tp\_speech$ presentation is produced by a speech on the speaker device and the $present\_tp\_expression$ presentation is returned by facial expressions on the screen.

$$(speech, speaker) = AFF(present\_tp\_speech)$$
$$(expression, screen) = AFF(present\_tp\_expression)$$

Event-B modelling

Event-B Method

The Event-B method (Abrial 2010) is a formal method based on first order logic and set theory, it is supported by the Rodin platform[1]. An Event-B model encodes a state transition system where the variables represent the state and the events represent the transitions. The Event-B modelling process is incremental, it starts from an abstract model of the system which evolves progressively to a concrete one by adding design details through successive refinement steps. The description of the Event-B model is associated to proof obligations ensuring the consistency of the model. These proof obligations are automatically generated and must be proved in order to ensure model correctness. An Event-B model is divided into two components (see Fig.3): CONTEXT which lists the static properties of the model and MACHINE which describes dynamic properties of the model (behavior). A Context is defined by a set of clauses

[1]http://www.event-b.org/install.html

CONTEXT $context_1$	**MACHINE** $machine_1$
EXTENDS $context_2$	**REFINES** $machine_2$
SETS ...	**SEES** $context_1$
CONSTANTS ...	**VARIABLES** ...
AXIOMS ...	**INVARIANTS** ...
THEOREMS ...	**THEOREMS** ...
END	**VARIANT** ...
	EVENTS ...
	END

Figure 3: The structure of an Event-B development

as follows: (1)EXTENDS declares the Context extended by the described Context. (2) SETS describes a set of abstract and enumerated types. (3) CONSTANTS represents the constants used by a model. (4) AXIOMS describes, in first order logic, the attribute properties defined in the CONSTANTS clause, types and constraints are described in this clause as well. (4) THEOREMS are logical expressions that can be deduced from the axioms. Similarly to Contexts, a Machine is defined by a set of clauses: (1) REFINES declares the Machine refined by the described Machine. (2) SEES declares the list of CONTEXTs imported by the described Machine. (3) VARIABLES for the declaration of state variables, refinement may introduce new variables in order to enrich the described system. (4) INVARIANTS describes VARIABLE properties, typing information, functional and safety properties are usually described in this clause. These properties shall remain true in the whole model. Invariants need to be preserved by events. They also express the gluing invariant required by each refinement. (5) THEOREMS is a set of logical expressions that can be deduced from the invariants. (6) VARIANT introduces a decreasing natural number for events termination. (7) EVENTS defines all the events (transitions) of a given model. Each event is characterized by its guard and is described by a body thanks to actions. Each Machine must contain an "*Initialisation*" event. The events occurring in an Event-B model affect the state described in VARIABLES clause.

Modelling approach

The Event-B development process we propose is based on the principle introduced in (Ait-Ameur et al. 2009). The authors propose to encode process algebra operators in Event-B, using an explicit variant to encode the events order and successive refinements. The left hand side part of a BNF rule is modelled by an abstract machine, and the right hand side part by a refinement of this abstract machine. The same principle is applied to the allocation model syntax rules. It gives rise to the following process: (1) development of the abstract machine related to PME; (2) refinement of the abstract machine to introduce elementary, iteration, complementary, redundant and choice operators as a decomposition. (3) refinement of the refined machine to allocate $(modality, media)$ pairs. The resulting model expresses the static and dynamic

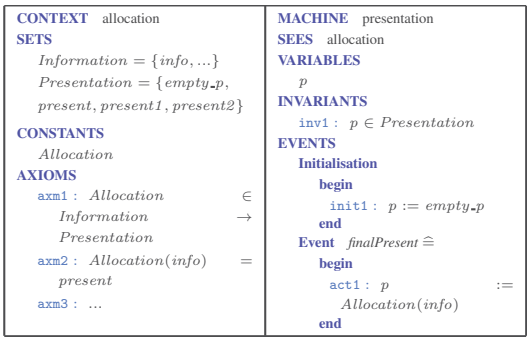

Figure 4: The abstract allocation Event-B model

properties defined by the generic allocation model. Thus, according to the Event-B development structure, the context component describes the static properties of the model and the machine component describes the dynamic behavior conforming to defined behavioral properties.

Event-B models

A set of generic Event-B models have been defined for each operator. Due to paper length limitation, we illustrate the Event-B models that formalize the $allocation$ step for the complementary operator, and the $affectation$ step.

Allocation model. The abstract allocation model describes the construction of the final multi-modal presentation. It contains the following components:
1. a CONTEXT $allocation$ defining the $Information$ and $Presentation$ sets (see Fig. 4). The $Presentation$ set contains: $empty\_P$ (empty presentation), $present$ (final presentation), $present1$ and $present2$ (elementary presentations). It also defines the $Allocation$ function that allocates presentations to information ($axm1$);
2. a MACHINE $presentation$ describing the construction of the p final presentation. This presentation is initialized to the $empty\_p$ value and computed by the $finalPresent$ event. Its final value corresponds to the presentation allocated to the $info$ information. (see Fig. 4). The refinement of the abstract model encodes the decomposition of the p presentation into a complementary pair of presentations ($p1$, $p2$). The obtained refinement contains:
1. a new CONTEXT $complementary$ (see Fig. 5), extending the $allocation$ CONTEXT by: (1) the complementary function $compl$ that assigns to each complementary presentation pair, the resulting presentation ($axm1$) (2) the valued definition of the $compl$ function ($axm2$, $axm3$, ...).
2. a refinement MACHINE $presentation\_compl$ (Fig. 5) refines the MACHINE $presentation$ by introducing the variables $p1$ and $p2$ that represent the two complementary elementary presentations that compose p, and two additional events: $presentation1$ and $presentation2$ to produce respectively $p1$ and $p2$. These events are followed by the $finalPresent$ event which refines the $finalPresent$ event of the $presentation$ MACHINE and produces the presentation p in terms of $p1$ and $p2$ presentations by the $compl$ function. We assume that $presentation1$, $presentation2$ events are triggered in parallel. The parallel scheduling is formalized by the introduction of the $varParallel1 + varParallel2$ variant according to (Ait-Ameur et al. 2009). Invariants $inv4$ and $inv5$ establish that the final presentation computed in the refinement is the same as the one computed in the abstract machine.

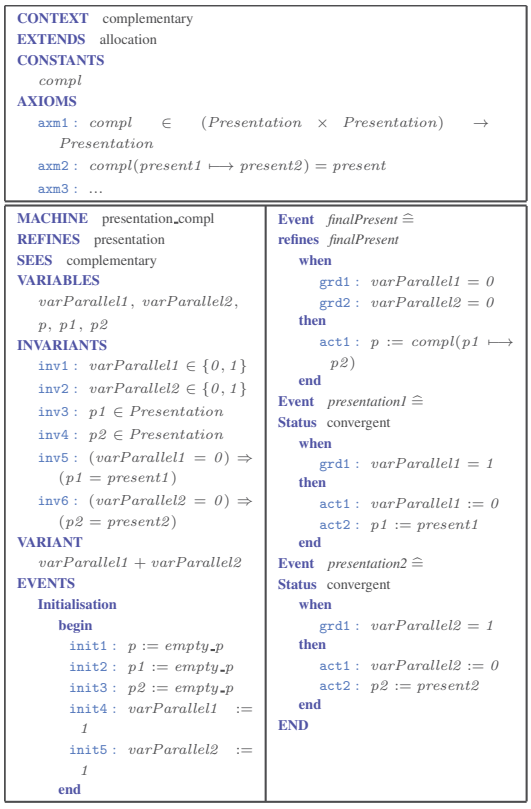

Figure 5: The complementary allocation Event-B refinement

Affectation model. The affectation model describes the affectation of a pair ($modality$, $media$) to each elementary multi-modal presentation unit produced by the elementary multi-modal presentations decomposition. The affectation Event-B model is composed of (see Fig. 6):
1. a new CONTEXT $affectation$ that defines: (1) the $Modality$ and $Media$ sets (2) the $affectation$ function assigning to each presentation unit, a pair ($modality$, $media$) ($axm1$).
2. a refinement MACHINE $presentation\_affect$ refines $presentation\_compl$. It affects for each presentation unit its corresponding pair ($modality$, $media$). The state of the $presentation\_affect$ MACHINE is described by the $item$ variable that formalizes the media and the modality used by the obtained multi-modal presentation ($act1$). The $affectation$ function is used to get the ($modality$, $media$) value associated to the p presentation in the $modality\_media$ event.

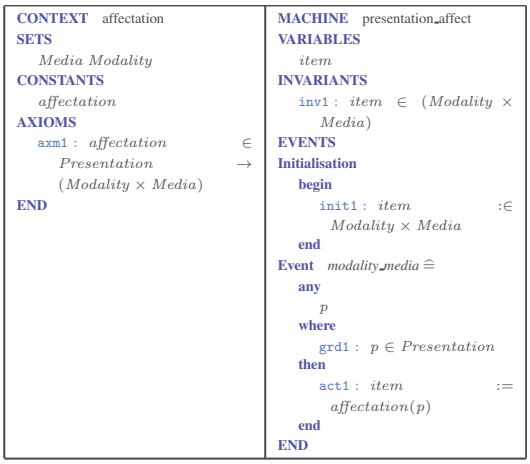

Figure 6: The affectation Event-B model

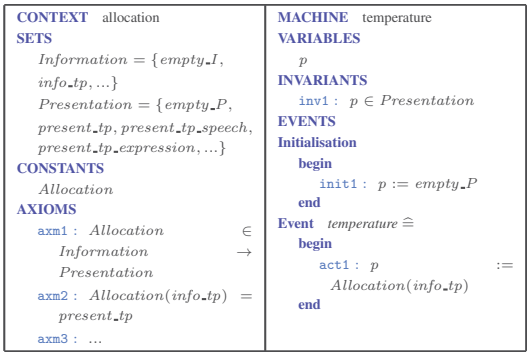

Figure 7: The temperature Event-B model

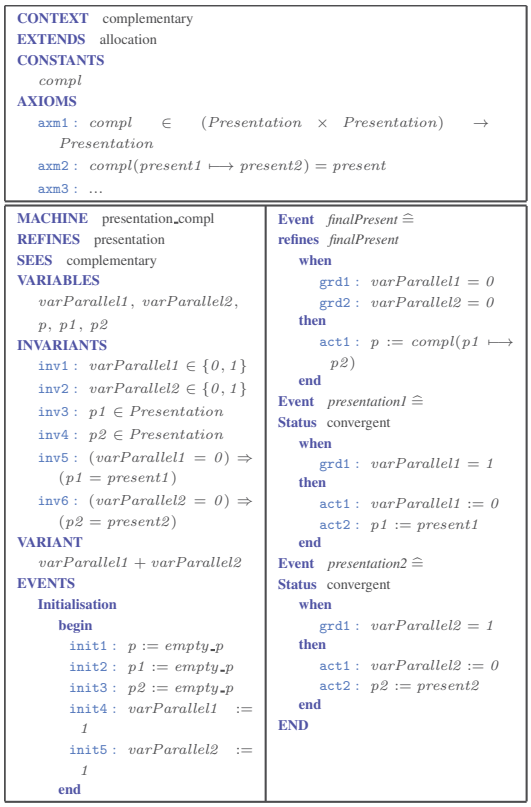

Figure 8: The decomposition Event-B model

Application to the defined case study

The application of the proposed approach to the case study previously introduced consists in enumerating the values of different sets and functions defined in the Event-B models of the previous section. Three successive models are obtained.

1. An abstract model (Fig. 7) builds the multi-modal presentation related to the information: "temperature over 50°C"; This model contains the CONTEXT $allocation$ defining $Information$ and $Presentation$ sets. $info\_tp$ formalizes the "temperature over 50°C" information and the $present\_tp$ the presentation to be build. The event $finalPresent$ of the $temperature$ MACHINE builds the final presentation $present\_tp$ allocated to $info\_tp$.

2. A first refinement (Fig. 8) defines the final presentation construction by using the complementary property between two elementary presentations. It contains the $complementary$ CONTEXT defining the $Compl$ function. This function decomposes $present\_tp$ into two complementary presentations: $present\_tp\_speech$ and $present\_tp\_expression$. This decomposition is performed in the $temperature\_compl$ MACHINE by the $temperature\_speech$ and $temperature\_expression$ events. These events build in parallel the $present\_tp\_speech$ and $present\_tp\_expression$ elementary presentations.

3. A second refinement (Fig. 9) refining the $temperature\_speech$ and $temperature\_expression$ events by affecting to $present\_tp\_speech$ and $present\_tp\_expression$, the ($modality$,$media$) pairs.

Properties verification

The proposed allocation models are generic. They make it possible to verify properties. First, deadlock freeness properties are expressed into the different machines by guard disjunction invariants ensuring that, at any time, the guard of at least one event is enabled. In addition, it's possible to verify other properties by introducing specific refinements. As example, a safety property, called collisions freeness is checked. It ensures that no collision may occur on a non-shareable media. A collision consists in producing two modalities in the same temporal window, on a non shareable media such as presenting, at the same time, a tone with a sentence produced by speech synthesis. In order to guarantee such property, it is necessary to supervise media allocation. Hence, collision freeness verification introduces a new refinement, successive to affectation. In this refinement, we need to know, at any time, the number of modalities allocated to the media. The collision freeness property is verified by introducing an invariant expressing that if the media is not

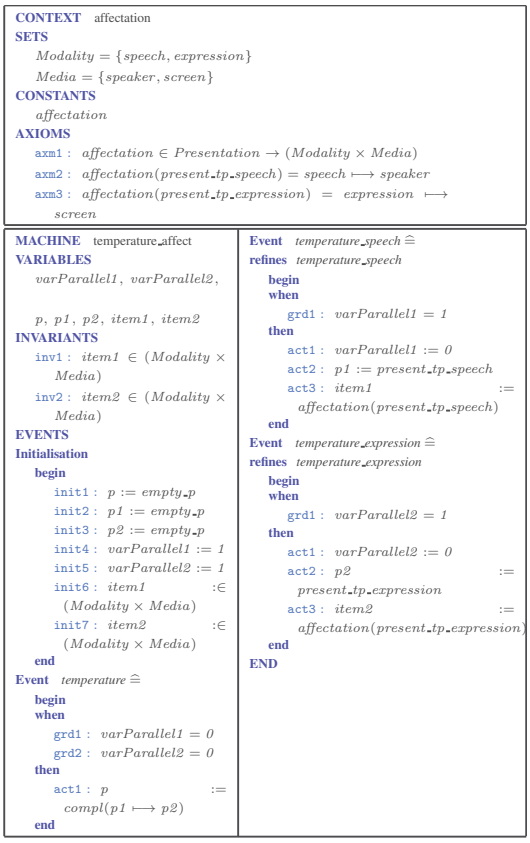

Figure 9: The affectation Event-B model

shareable, then the cardinality of using this media must be less than 1 (the media can not be allocated to more than one modality). The refinement formalizing the collision freeness property is not presented in this paper.

Conclusion

This paper deals with the modelling of output multi-modal human-machine interfaces. It extends previous work proposing a generic formal model. It proposes an Event-B based formal modelling of the modality/media allocation step. On the one hand, it allows the interface designer to formally describe the allocation process (chosen combination of modalities and media) and secondly, it enables properties validation by the theorem proving mechanism associated to Event-B modelling process. Some safety and liveness properties validation have been performed, and it would be interesting to carry out the validation of other properties e.g. usability properties to show that our model is generic enough.

References

Abrial, J.-R. 2010. *Modeling in Event-B: System and Software Engineering*. Cambridge University Press.

Ait-Ameur, Y.; Baron, M.; Kamel, N.; and Mota, J.-M. 2009. Encoding a process algebra using the Event B method. *International Journal on Software Tools for Technology Transfer*.

Bordegoni, M.; Faconti, G.; Maybury, M.; Rist, T.; S. Ruggieri, P. T.; and Wilson, M. 1997. A standard reference model for intelligent multimedia presentation systems. *Computer Standards and Interfaces*.

Bouchet, J.; Madani, L.; Nigay, L.; Oriat, C.; and Parissis, I. 2008. Ingénierie systèmes interactifs. chapter Test Formel de systèmes interactifs multimodaux.

Kamel, N., and Ait-Ameur, Y. 2007. Formal model for usability properties in multimodal HCI. In *(MAPS07), at IEEE International Conference On Pervasive Services (ICPS07)*.

Mohand-Oussaid, L.; Ait-Ameur, Y.; and Ahmed-Nacer, M. 2009. A generic formal model for fission of modalities in output multi-modal interactive systems. In *VECoS '2009*.

Mohand-Oussaid, L.; Ait-Sadoune, I.; and Ait-Ameur, Y. 2011. Modeling information fission in output multi-modal interactive systems using event b. In *(MEDI) 2011*.

Mohand-Oussaid, L.; Kamel, N.; Ait-Sadoune, I.; Ait-Ameur, Y.; and Ahmed-Nacer, M. *Human computer interaction in transport*. chapter A formal framework for design and validation of multimodal interactive systems in transport domain.

Navarre, D.; Palanque, P.; Bastide, R.; Schyn, A.; Winckler, M.; Nedel, L. P.; and Freitas, C. M. 2005. Une description formelle des techniques d'interaction multimodales pour immersive virtual reality applications. In *IFIP TC13*.

Palanque, P., and Schyn, A. 2003. A Model-based for Engineering Multimodal Interactive Systems. In *(Interact'2003)*.

Rousseau, C.; Bellik, Y.; and Vernier, F. 2005. WWHT: Un modèle conceptuel pour la présentation multimodale d'information. In *IHM2005*, 59–66.