Journal of Web Engineering, Vol. 4, No.1 (2005) 038-056

FUNDAMENTALS OF EXCEPTION HANDLING WITHIN WORKFLOW-BASED WEB APPLICATIONS[a]

MARCO BRAMBILLA CHRISTINA TZIVISKOU
Dipartimento di Elettronica e Informazione, Politecnico di Milano
Via Ponzio 34/5, 20133 Milano, Italy
+39 02 2399 3408
mbrambil@elet.polimi.it tzivisko@elet.polimi.it

Received October 5, 2004
Revised November 20, 2004

As the Web becomes a platform for implementing B2B applications, the need arises of extending Web conceptual modeling from data-centric applications to data- and process-centric applications. New primitives must be put in place to implement workflows describing business processes. In this context, new problems about process safety arise, due to the loose control on Web clients. Indeed, user behavior can generate dangerous incoherencies for the execution of processes. This paper presents a proposal of workflow-enabling primitives for Web applications, and a high level approach to the management of exceptions that occurs during execution of processes. We present a classification of exceptions that can occur inside workflow-based Web applications, and recovery policies to retrieve coherent status and data after an exception. An implementation experience is briefly presented too.

Key words: Web site design, web site management, workflow, exception handling

1 Introduction

In recent years, the Web is more and more being used as the implementation platform for B2B applications, whose goal is not only the navigation of content, but also supporting business processes, content management, value-added services and so on. Conceptual modeling expertise from other fields (database, object-orientated programming, and hypermedia applications) has been widely recognized as valid starting point for defining conceptual aids for Web application development too [8]. The first generation of conceptual models for the Web [1, 2, 6, 7] essentially focus on capturing the structure of data to be published, and the navigation primitives, represented by such concepts as pages, content nodes, and links.

To cover business processes support, a second generation of conceptual models is required. These new models should cope with process and workflow modeling, support Web service interaction, and integrate data-centric and process-centric modeling primitives into a mix suited to the development of advanced B2B Web applications. In this context, it is important to address the critical cases that can occur in the enactment of business processes on a Web-based platform.

This paper presents an extension to a first-generation Web modeling language [6, 7, 15] to support the specification, design and implementation of B2B applications, and offers an high-level analysis of critic aspects and exception management issues within Web applications exploiting business processes.

[a] This research is part of the WebSI (Web Service Integration) project, funded by the EC in the Fifth Framework.

Exceptions that can happen in a Web based application have peculiar characteristics with respect to traditional workflow applications. This is due to two main aspects: (i) interaction options provided by browser-based interfaces are very powerful, but they are more oriented to free navigation than strict processes adherence (e.g., user is enabled to jump back and forth on navigated pages, thus introducing dangerous repetition of process activities); (ii) user cannot be forced to perform any action or task (e.g., he can stand on a page for long time, or even close the browser and disconnect at any time).

Our approach is lightweight: we are interested in extending Web modeling to cope with process and exception modeling, not to adapt workflow management systems (WfMS) to the Web. About exceptions, we aim at defining a modeling paradigm for critical cases, not to build transactional systems or low-level exception handling mechanisms.

The paper is organized as follows: Section 2 reviews the related work; Section 3 briefly outlines the main concepts about workflow and Web application modeling, providing an overview of standards and notations in the Web engineering and workflow modeling fields; Section 4 introduces the study of critical situations that can occur in the execution of processes on the Web; Section 5 defines the concept of exception and a possible categorization; Section 6 presents our approach to management and recovery of exceptional situations within process execution; this includes the refined description of activities, and the metadata and hypertext primitives for managing exceptions. Section 7 describes the recovery policies for exceptional situation and provides a classification of the policies with respect to different dimensions. Finally, Section 8 reviews implementation experience and Section 9 draws some conclusions and presents our ongoing and future work.

2 Related work

In the Web Engineering field, exception handling has always been neglected by high level modeling approaches. This is probably due to the fact that research in this field is not yet mature; therefore it needs to put in place the foundation of the scientific approach before focusing on exceptional situations.

In other fields (database, software engineering, and so on), many works have addressed the problem of exception interception and compensation. They mainly studied transactional properties for activities, which is not in our scope. However, some works deals with weaker properties. For example, [9] is based on the concept of spheres, to make use of only those transactional properties that are actually needed; [14] is one of the first works that address the problem in the Web context.

Recently, some initial studies on exception handling specifically addressing the Web environment have been developed. In [14], Miller et al. study the problem in the Web context, but they make only an exception classification without proposing further handling mechanisms. A few other existing methodologies address exception handling only with respect to transactional properties within activities. Among them, we can distinguish OO-H and UWE.

OO-H [4] is a partially object-oriented approach for modeling Web applications integrated with business processes. The methodology steps, clearly separated, are defined for the design of the conceptual, process, navigation and presentation model. For process modeling, the approach provides a set of mapping rules that transform the underlying process definition in the user navigation model; therefore the navigation structure is determined by the workflow model; the user navigation of links corresponds to the process flow.

Also in UWE [10], an object-oriented method, the process design is specified by separated models. Alike OO-H, the navigation model is extended with process primitives that transfer the user inside and outside the boundaries of the process flow. In both methods, the data model is implemented by means of UML class diagrams and is integrated with process elements. There is no explicit process metadata: tasks are directly modeled as class methods and process activities are performed by invoking these methods. Consequently, process monitoring and management results a difficult task. In our work, the data model is based on the Entity-Relationship schema, and is extended with workflow and exceptions metadata, therefore it explicitly provides information about the process structure and flow.

In OOHDM [17], the content and navigation models are extended with activity entities and activity nodes respectively, represented by UML primitives. Furthermore, the process execution occurs within a navigational context that specifies the access rules for the corresponding process. Possible problems arising from user navigation within process execution are controlled by the navigational context, by means of special links and other mechanisms.

In WSDM [16], the process design is driven by the user requirements and is based on the ConcurTaskTrees notation. The actual process modeling is specified at the conceptual design. In particular, during the first phase that is the Task Modeling, the tasks hierarchy is defined, the temporal conditions among tasks are expressed by *operators*, and the information and / or functionality required by each task are modeled by *object chunks*. During the second phase of the conceptual design, the Navigational Design, the user navigation structure is generated by means of *components* and *project logic links* in order to perform the modeled tasks. Also in this approach, exception handling is achieved just at level of transactions defined for combined tasks.

On the other hand, our exception handling proposal aims at providing a more complete solution to the problem, by providing a classification of exceptions, together with handling policies and implementation guidelines. Our exception handling approach is based on a high-level modeling language called WebML [6, 7, 15], that has been recently extended for supporting workflows [3] and Web services. With respect to [3], the specific contribution of this paper is the introduction of the exception handling problem in the picture. Indeed, our previous works only described the workflow modeling within hypertext, without considering failures and exceptions.

Since we do not aim at adapting workflow management systems to the Web, traditional WfMS applications like Microsoft MQSeries [11], Oracle Workflow [12], and many others are not considered as direct competitors of our approach, even if some of them provide some Web-oriented facilities.

2 Conceptual modeling of Web applications and workflows

Conceptual design consists in high-level, platform-independent specification of the application, which can be used to drive the subsequent implementation phase. In this section we focus on two aspects of conceptual design: (i) Web application design, briefly describing the WebML model, which will be used in the sequel to describe our proposals; (ii) Workflow modeling concepts and primitives.

It is important to point out that, although the paper uses the WebML notation to describe our contribution, the proposed approach is independent from the specific language or notation that is adopted. Our approach to conceptual design relies on the following guidelines: an Entity-Relationship diagram models the data stored, manipulated, and exchanged by the application actors, plus the metadata required for the management of the business processes; *process diagrams* are treated as a higher-level specification and are used to derive a set of hypertext models that "realizes" them. These

hypertext models belong to the site views of the user groups involved in the process and must offer them the interface needed for performing their activities.

2.1 Process modeling

For specifying processes, we adopt the terminology defined by the Workflow Management Coalition [20], which provides a workflow model based on the concepts of Process (the description of the supported workflow), Case (a process instance), Activity (the elementary unit of work composing a process), Activity instance (an instantiation of an activity within a case), Actor (a user role intervening in the process), and Constraint (logical precedence among activities and rules enabling activities execution). Processes can be internally structured using a variety of constructs: sequences of activities, AND-splits (a single thread of control splits into two or more independent threads), AND-joins (blocking convergence point of two or more parallel activities), OR-splits (point in which one among multiple alternative branches is taken), OR-joins (non-blocking convergence point), iterations for repeating the execution of one or more activities, pre- and post-conditions (entry and exit criteria to and from a particular activity).

For the pictorial representation of workflow, we adhere to the Business Process Management Notation [BPMN] specification, which provides a wide set of symbols for visual representation of processes, according to the [BPML] standard. The BPMN notation allows to represent all the basic concepts defined by WfMC, and provides further expressive power through more powerful conditional gateways, event and exception management, free combination of split/join points, and other minor features. Fig. 1 briefly summarizes the main visual items provided by BPMN.

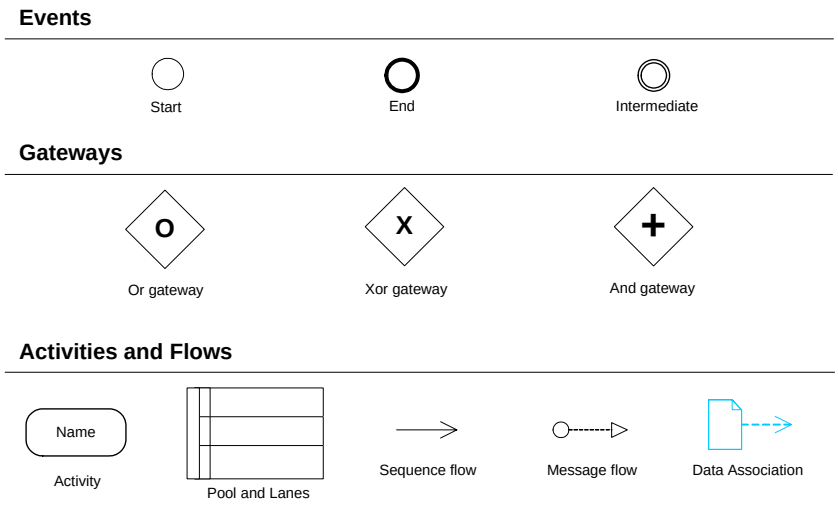

Fig. 1. BPMN workflow notation

Events are happenings that occur during the process execution, producing an impact. Events are categorized by type in BPMN. Types (not shown in the picture) include message, timer, rule, and other events. Appropriate symbols for event typing can be put into the circle representing the event.

Gateways are elements that control the flow of the process. They can be used either as decision, splitting, and merging and synchronization points. Various logical behaviors are allowed for the gateways.

Activities correspond to the WfMC definition, and moreover they can express several different behaviours (cycling, compensation, internal subprocess structuring, event catching, and so on).

The flow of the process is described by means of arrows, that can represent either the actual execution flow, or the flow of exchanged messages, or the association of data objects to activities. The latter is not meant to influence the execution of the application; it simply provides additional information about the used data.

Activities are grouped based into pools based on the participant that is in charge of the activity enactment. Typically a participant is identified as an organization that plays some roles. In our case, we will consider a participant as a peer of a distributed process. Pool lanes will be used to distinguish different user types that interact with the specific peer.

Fig. 2 exemplifies a workflow specifying the process of online purchase, payment and delivery of goods. The customer can choose the products to purchase, then submits his payment information. At this point, two parallel tasks are executed by the seller employees: the warehouse manager registers the shipping of the order, and a secretary prepares a bill to be sent to the customer.

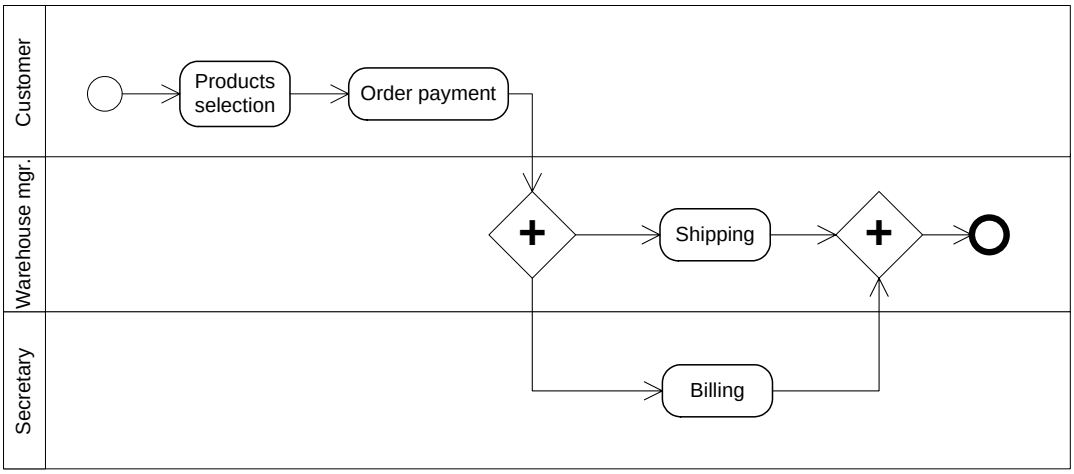

Fig. 2. Workflow diagram of the refunding request process

2.2 Hypertext modeling

For hypertext modeling, we use the WebML notation[6, 7, 18], a conceptual language for specifying Web applications developed on top of database content described by a E-R diagram. A WebML schema consists of one or more *site views*, expressing the Web interfaces that allow the different user roles to browse or manipulate the data specified in the underlying E-R schema. A *site view* contains *pages*, possibly clustered in *areas*, typically representing independent sections of the site. Pages enclose *content units*, representing atomic pieces of information to be published (e.g., indexes listing items from which the user may select a particular object, details of a single object, entry forms, and so on); content units may have a *selector*, which is a predicate identifying the entity instances to be

extracted from the underlying database and displayed by the unit. Pages and units can be connected through *links* of different types to express all possible navigation.

Besides content publishing, WebML allows specifying *operations*, like the filling of a shopping cart or the update of content. Basic data update operations are: the creation, modification and deletion of instances of an entity, or the creation and deletion of instances of a relationship. Operations do not display data and are placed outside of pages; user-defined operations can be specified (e.g., e-mail sending, e-payment, ...), and operation chains are allowed too.

Fig. 3 shows a simplified version of the two areas of the Customer site view of the e-commerce site example, whose workflow have been illustrated in Fig. 2: the *Products* area allows guests to browse products, by selecting in the *Home* page the product group from an index (*ProductGroups*). Once a group is selected, all the products of that group are shown in page *Products*. The *Mailing List Subscription* area allows the user to subscribe to a mailing list through a form. The submitted data are used to modify the profile of the User, by means of a modify operation called *Modify Subscr*, which updates the instance of entity *User* currently logged.

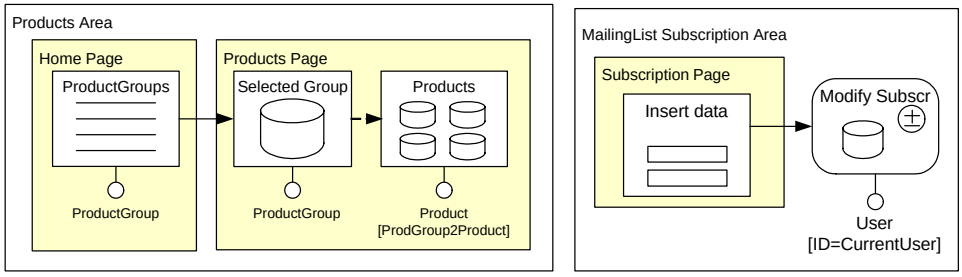

Fig. 3. WebML site view diagram featuring areas, pages, content units, and operations.

2.3 Extending hypertext modeling to capture processes

In the specification of a Web application supporting business processes [3], the data model, which is normally used to describe the domain objects, is extended with user-related and workflow-related data, and the hypertext model is enriched by a set of primitives enabling workflow dependent content of pages and navigation.

Process metadata. Data modeling is extended with the metadata used to represent the runtime evolution of processes as shown in Fig. 4. The schema includes entities representing the elements of a WfMC process model, and relationships expressing the semantic connections between the process elements.

Entity *Process* is associated with entity *ActivityType*, to represent the classes of activities that can be executed in a process. Entity *Case* denotes an instance of a process, whose status can be: initiated, active, or completed. Entity *ActivityInstance* denotes the occurrence of an activity, whose current status can be: inactive, active and completed. Entities *User* and *Group* represent the workflow actors, as individual users organized within groups (or roles). A user may belong to different groups, and one of these groups is set as his default group, to facilitate access control when the user logs in. Activities are "assigned to" user groups: this means that users of that group can perform the activity. Instead, concrete activity instances are "assigned to" individual users, who actually perform them. If needed, the model can be enriched at will with new relationships to represent more complex assignment rules.

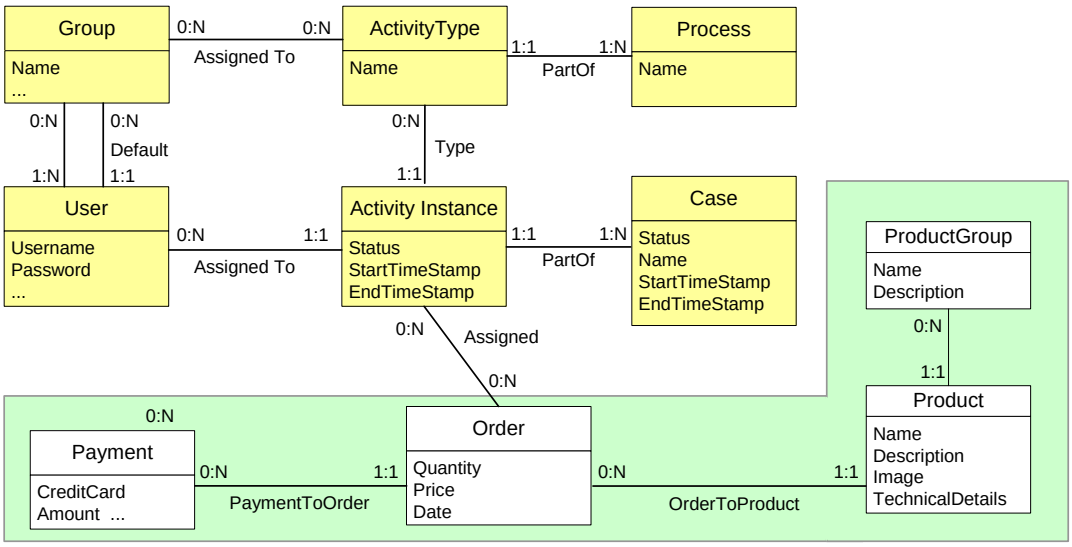

Fig. 4. Data model incorporating workflow concepts

Application data is described by a usual E-R model representing information involved in the current application. In our example, as depicted in the boxed part of Fig. 4, we model a catalog (in which each *Product* belongs to a *ProductGroup*), the *Orders* that the user submits and the *Payment* details. Orders are *assigned* to Activity Instances in which will be processed, whilst Payments are connected to the Activity Instances in which they are created. These relationships associate metadata concepts to application information. In general, the designer can specify an arbitrary number of relationships between the application data and the workflow data, which may be required to connect the activities to the data items they use. Note that minimum cardinality of these relationships is typically 0, since in most cases each activity instance is not associated to all the application data, but only to a very small set of objects.

Workflow hypertext primitives. In order to enact the process, some workflow-specialized hypertext primitives are also necessary to design interfaces capable of producing and consuming such metadata. At this purpose, a few additional primitives are introduced in WebML for updating process data as a result of activity execution, for accessing the data associated with a specific activity instance, and for expressing the assignment of data objects to an activity instance eventually to be executed.

The portion of hypertext devoted to the execution of an activity must be enclosed between the two workflow-related operations shown in Fig. 5 (a): *start activity* and *end activity*. These operations are triggered respectively by incoming and outgoing links of the activity and have the side effect of updating the workflow data. Specifically, starting an activity implies creating an activity instance, recording the activity instance activation timestamp, connecting the activity instance to the current case (relationship PartOf), to the current user (relationship AssignedTo), and to the proper activity type, and setting the status of the activity instance to "active". Symmetrically, ending an activity implies setting the status to "completed" and recording the timestamp.

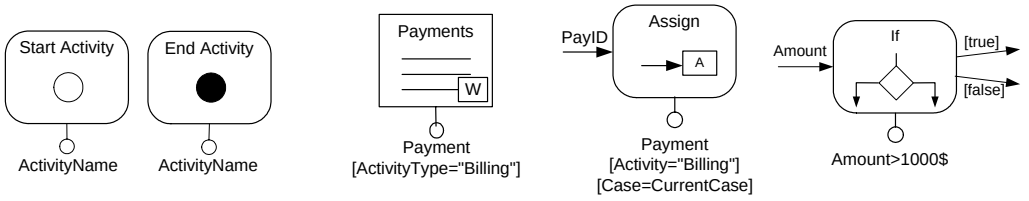

Fig. 5. *Start Activity* and *End Activity* operations (a); workflow-aware content unit notation(b); graphical notation of the *Assign* operation (c) and of the conditional operation (d)

The *Start Activity* operation can also be marked as the *starting case activity*, when the activity to start is the first one of the entire process; dually, the *End Activity* operation can be tagged as the *end of the case*, thus recording the general status of the process.

Workflow-aware content units can be used for retrieving the data objects related to a particular activity. These units are like the regular WebML content unit but are tagged with a "W" symbol denoting a simplified syntax for their selector, which shortens the expression of predicates involving both application data and workflow data. For example, Fig. 5 (b) shows a workflow-aware index unit that retrieves all the instances of entity *Payment* that have been assigned to an activity of type “Billing”.

The *assign operation* is a WebML operation unit that connects application object(s) to an activity instance, for which an activity type, a case and possibly a user are specified. Fig. 5 (c) shows the graphical representation of the assign operation, which assigns a *Payment* to the activity called "Billing" for the current process case.

The navigation of a hypertext may need to be conditioned by the status of activities, to reflect the constraints imposed by the workflow. Two dedicated operations called *if* (see Fig. 5 (d)) and *switch* operations allow conditional navigation, performing the necessary status tests and deciding the destination of a navigable link.

Mapping rules have been defined from WfMC-based workflow description to WebML hypertexts enhanced with workflow primitives [3].

3 Critical Situations and Exception sources

Within the execution of a process, exceptional situations can occur, due either to user behavior or to system failures. We define a critical situation as an incorrect browsing behavior of the user (*user-generated exceptions*) or a technical failure of the system (*system failures*).

3.1 User-generated exceptions

This section presents the critical situations that can arise from wrong browsing behavior. For Web context, this problem is much more relevant than for traditional applications. The most evident examples are *back* and *forward* buttons of a Web browser, that allow the user to explore the hypertext of the Web application in a free way, while a workflow scenario has usually a strictly forced execution/ navigation structure, and its steps must be executed in the proper order. Moreover, the user is able to jump without restrictions from an application to another. Back and forward buttons let the

user go outside the pages of an activity still active or move back to a completed activity and try to resume its execution. With respect to workflow activities, improper browsing can be of three types:

- *improper inbound browsing:* the user gets into a workflow activity without executing the *Start activity* operation, for example by clicking repeatedly on the *back* browser button, until a previously executed activity is reached;
- *improper outbound browsing:* the user, during the execution of an activity, follows a *wrong* navigational path, exiting the activity without passing by the *End activity* operation. In this case the user leaves the pages of the current activity, either by pressing repeatedly the back browser button or by following a landmark link (i.e., a link which is always clickable within the whole Web application). In this way, the user can potentially start an arbitrary number of activities, since he can try to start a new activity beside the current one. Moreover, the user left an activity in status Active, which cannot proceed, and thus remains halted;
- *improper internal browsing:* the user, during the execution of an activity, presses the back button of the browser one or more times reaching a previous page of the same activity, and then clicks on a link, trying to repeat part of the activity. In this way, the user is in a page that is different from the current step of the activity, since the page from which the user resumes the browsing is different from the last page requested to the server;
- *wait:* the user does not request a page to the server for a given amount of time, after which a timeout expires and the user session ends up. A *Session End* exception is generated, and this behavior collapses in a system failure.

3.2 System failures

System failures can occur both at client and at server side. Client-side failures are problems that are generated by system breakdown, which is either a client crash or a network failure. We do not consider server-side failures, since this problem for Web-based workflows can be addressed in the same way of traditional workflow systems, and several recovery theories and techniques already exist for this context (e.g., rules based on active rules [5]). System failures result in a Session end exception at server-side. To discover client-side failures, HTTP session is a standard technique employed in Web applications. After a session has been established, a network failure or a client failure will result either in the client not performing a request to the server for a given amount of time, or in the server being unable to send the response back to the client. When the server recognizes that the client is no more reachable, it will end up the user session: client-side failures can be captured at application level by generating a Session End exception. In this sense, client failure and network failure will be indistinguishable and will be collectively denoted also as Crash situations.

After a crash situation the activity instance executed by the user remains in Active status, but is not completed. This means that the activity execution cannot proceed, since the user lost his session, and if he tries to login he can only see the activities that are in Inactive status (ready to be executed). Typically he is not allowed to perform activities potentially in execution (i.e., in Active status).

If the activity instance is not recovered, the whole process case will possibly be stopped, if there are other activities waiting for the completion of the crashed one.

A thrown Session End exception will help to track the crash for later recovery.

3.3 Inconsistencies

Data and process inconsistencies can arise from system failure and incorrect browsing behavior. Each of them will be addressed with a different approach:

- *activity/process halt:* one or more activities (and the processes they belong to) get halted and cannot be resumed or concluded by the user. These problems are detected after they take place and are recovered by means of appropriate policies;
- *inconsistent database:* one or more database tuples are created or destroyed in an unexpected way, resulting in an inconsistent database and workflow application; these problems are caused by incorrect browsing behavior, and will be handled in a preventive way, by detecting the user faults and generating an exception before they result in a failure.

4 Exception definition and categorization

As we have seen in previous sections, if a critical situation occurs, the workflow application might be in an inconsistent state due to the presence of a halted activity, i.e. an activity in status Active that cannot proceed. The need arises to recover the halted activity to bring the workflow application back to a correct state and let the process execution proceed. To address the problem, we define the concepts of exception and recovery policy.

To manage critical situations and to prevent/recover inconsistencies, we introduce the concept of exception. An *exception* is an event that is thrown by the system, as a consequence of a critical situation that is occurred.

An exception is either synchronous, if it is thrown after a page request, or asynchronous, if it is not tied to a page request but can occur independently. In case of synchronous exceptions, the user navigation can be immediately affected since the server can decide to provide the user with a different page depending on the caught exception. On the other hand, the only asynchronous exception that we will consider is Session End. It cannot influence immediately the user browsing, since he already disconnected from the application (his session is no more valid). Table 1 resumes the characteristics of exception types.

Table 1. Types and properties of the exceptions

Exception Type	Session Status	Addressed Problem
Asynchronous	Inactive	Technical Failure Incorrect Browsing Behavior
Synchronous	Active	Incorrect Browsing Behavior

Exceptions to be managed in order to guarantee the correctness of workflow-based Web applications are the following:

- *Session End:* the user disconnected the client, or a failure happened on the network or at client-side. These events are undistinguishable from server side;
- *Activity Already Active:* the user is trying to start an activity when there is another activity already active in his session;
- *Wrong Starting Page:* inside an activity, an action has been performed in a page that is not the last one that the user has visited;
- *Action By completed Activity:* an action has been performed within an activity that has been already closed.

5 Primitives for exception handling

To study in a simple and effective way the exception handling problem, we define some new concepts that refine the description of the structure of activities and extend the data model and the hypertext model of the Web application.

5.1 Fine grained description of activities

We call *step* a hypertext page belonging to an activity. Steps are univocally numbered within an activity. Between two subsequent steps there can be a chain of operations, which is not relevant for our purposes. Indeed, since we do not consider server-side failures, a chain of operations can be seen as an atomic element that never fails (server-side failure is addressed by standard WebML mechanisms, like KO links [7]).

We define the *current step* of an active activity as the last page that the server has generated after a request by the client. This information is stored into the *CurrentStep* entity of the workflow metadata schema (see Fig. 7). Within a process case, it is always possible to retrieve the currently active activities, and for each of them the current step. The current step has 2 important properties:

(i) it is always uniquely defined for an active activity;
(ii) it gives us a correct idea of the progress of the activity.

It is important to notice that if the client uses the back and forward buttons of the browser, the current step of the activity does not change, since the client does not make any request to the server. Moreover, by clicking the *back* button we do not roll back the operations between consecutive steps, we just reload an old page at client side.

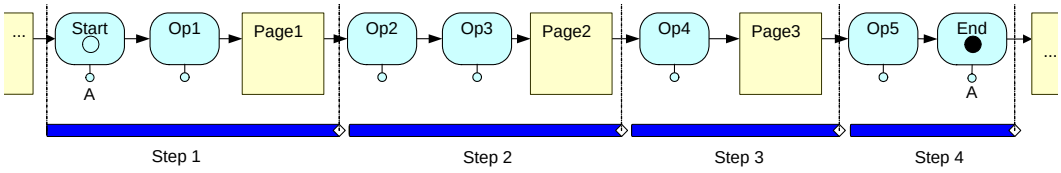

Fig. 6. Pictorial example of steps within an activity

The current step is fundamental to implement exception management mechanisms aiming at resuming the activity from the point it was abandoned. More details on this type of behavior is provided in the sequel of the paper. Fig. 6 visually represents the concept of current step: once the activity is started, CurrentStep is set to 1. Step 2 starts once the user clicks on the outgoing link in Page1and the corresponding HTTP request reaches the server. Indeed, at this point we are guaranteed that the server will perform all the operation chain (Op2 and Op3 in the example) and will generate Page2 for the client. Since CurrentStep is set to "2" for this period, I'm granted that possible exceptions and subsequent resume policies from CurrentStep will bring the user to the correct page (Page2), avoiding the repetition of already performed operations. The same applies to the other steps. Notice that Step 4 has no associated page. This means that, if CurrentStep is 4, the activity is actually granted to be concluded, since the link in page 3 has been clicked, the request has reached the server, and the operation chain has been completely executed.

5.2 Extending the data model to capture exceptions

The exceptions that take place in a Web application supporting business processes are tracked through a new metadata model, the *exception data model* shown in Fig. 7. This model includes metadata for supporting exception handling information and extends the workflow metadata model presented in Section 2.3. It represents generated exceptions, associated with the execution of process activities. The schema includes entities representing exception elements, and relationships expressing connections with the affected activities. The exception data model is.

The following new elements (represented in bold face in Fig. 7) have been added:

- *Created* relationship: connects the Activity Instance to the application data object that is managed within the workflow. It keeps track of the activity during which a specific object has been created, at the purpose of allowing possible removal of the object if the activity will be canceled. Further details on the use of the *Created* relationship will be provided in Section 5.3, when describing the Reject policy.
- *CurrentStep* entity: is needed for supporting some recovery policies for exceptions. It saves the current step of the exception, as described in Section 5.1. Its use will be explained in Section 5.3, when describing the Resume policy. The *Current* relationship associates the CurrentStep to the relationship it refers to.
- *Exception Instance* entity: represents the individual instantiations of an exception that occurs for a specific activity, described by the timestamp and the status of the exception. The status can be: active(i.e., the exception has occurred and has not been addressed yet), resolving (i.e., the exception is currently being solved by a predefined policy or a compensation chain), and resolved(i.e., the exception has been solved by some exception handling mechanism). The update of the status is delegated to each exception handling mechanism. In general, an exception may involve several activities at time. An exception instance is generated for every activity instance affected by the exception, and is connected to the activity it affects. The above connection is expressed in the schema by the *Affects* relationship.
- *Exception Type* entity: categorizes the occurred exceptions depending on their type, as defined in Section 4. The possible types are: Session End, Activity Already Active, Wrong Starting Page, and Action By completed Activity. Each exception that is instantiated is connected to its type through the *Type* relationship.

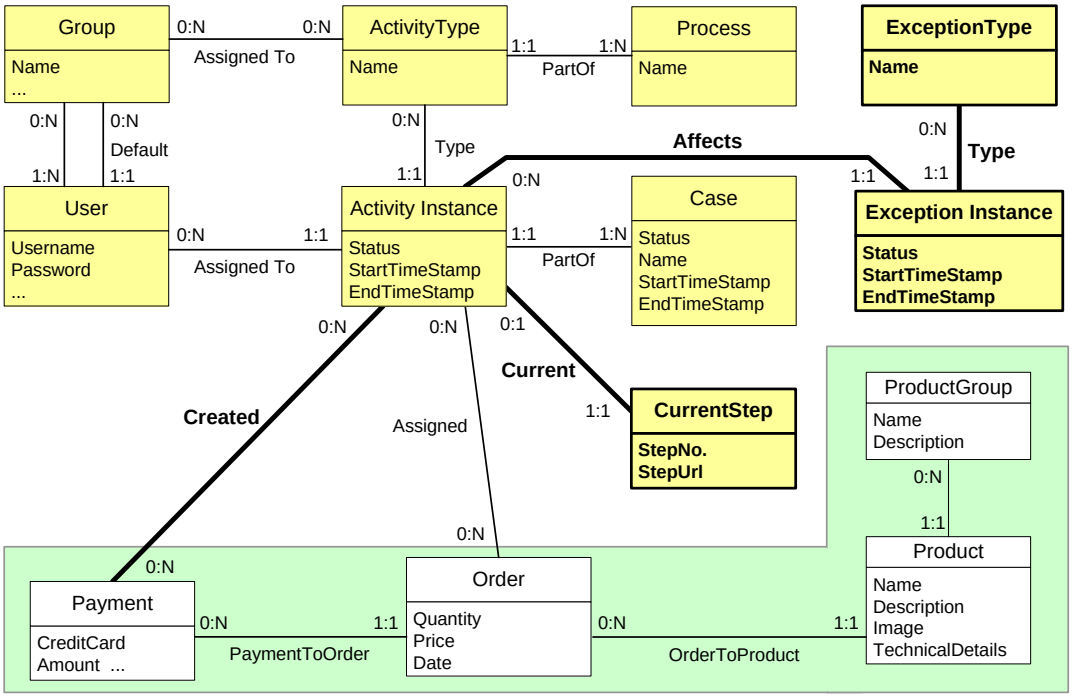

Fig. 7. Data model incorporating workflow concepts and exception information

The information stored within the Exception Data Model will be used and updated at runtime by the exception handling mechanisms that will be described in section 6.

5.3 Extending the hypertext model to capture exceptions

In order to handle generated exceptions, existing and new hypertext primitives are used to retrieve and modify the exceptions metadata denoted by the entities of the exceptions data schema. At this purpose, the insertion and update of exception data is achieved with the use of standard WebML operation units, like Create, Modify and Delete units. Moreover, we define two additional units for managing exceptions: *Exception-aware index unit* and *Catch event unit.*

For easily accessing the data associated with a specific exception instance, a new content unit is introduced. *Exception-aware index unit* can be used for retrieving the activity objects affected by a specific exception and displaying the available recovery policies. These units are tagged with an "E" symbol that denote a simplified syntax for their selector condition, which shortens the expression of predicates involving workflow and exception data (as in the workflow-aware index unit). The exception-aware index unit allows displaying the list of activities affected by a specific Exception Type, that has a specific Status, and involving a specific User. Some usage examples are shown in Fig. **8**: the exception-aware index unit is used to (a) retrieve the activities affected by a not yet solved Session End exception that are related to a specific user; (b) retrieve activities of type “Payment” affected by any type of exception that are related to a specific user and not solved yet; (c) retrieve activities not related to the current user that are currently addressed by an exception handling mechanism.

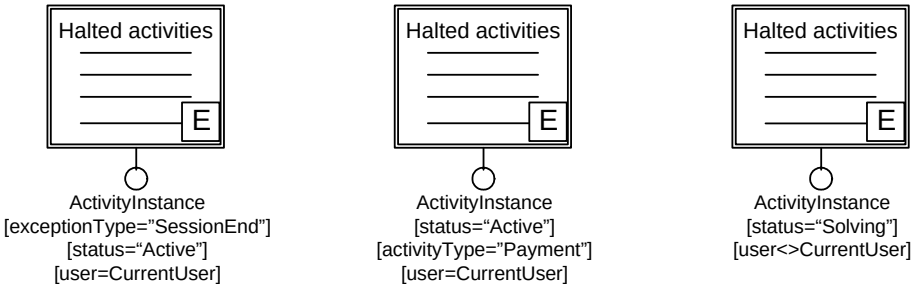

Fig. 8. Exception-aware index unit examples

The *CatchEvent* unit is a sort of placeholder for exception catching. It is provided with the following parameters: ExceptionType, ActivityType, ActivityInstance, ExceptionInstance.

ActivityType and ExceptionType are specified at design time and define the situation in which the compensation chain is triggered. ActivityType is a mandatory property that must be set for the unit, whilst ExceptionType can be omitted if the designer wants to capture any type of exception with the same unit. ActivityInstance and ExceptionInstance are runtime parameters, whose values are available to operations of the chain, for retrieving further related data.

This unit has the role of triggering the execution of the following operations once the exception is raised. Note that, since triggering and execution of these operation chains is completely automatic, no pages involving user interaction are allowed. In particular, this unit will be used for implementing customer-defined compensation chains, as described in Section 6.4. Fig. **9** shows some usage examples of the CatchEvent unit: the first unit captures all the possible exception that may arise within an instance of PaymentActivity; the second example captures all the exceptions of type "AlreadyStartedActivity" that may arise within an instance of PaymentActivity.

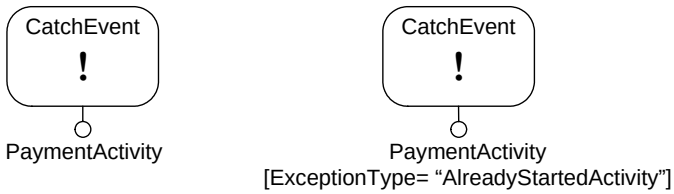

Fig. 9. CatchEvent unit examples

6 Recovery policies

We define a *recovery policy* (for a halted activity) as a collection of operations that we perform on the activity and on the related data in order to bring the workflow application to a correct state and to let the process proceed. Policy behavior can be very different, but they all are asked to perform some basic tasks in order to maintain the exception metadata model. In particular, they must update the Status of the Exception Instance they are managing, by setting its value to *Solving* at the beginning of the policy and then to *Solved* at the end of the policy. Other actions may vary depending on the meaning and on the purpose of the policy.

6.1 Policy classification

Policies can be classified with respect to three orthogonal dimensions:

- **policy direction**, that considers the way in which a coherent status of the process is reached: the policy can try to recover a correct status that was previously visited by the workflow application (*backward policy*), or can try to move to a new correct status that was not previously visited by the workflow application (*forward policy*).
- **policy definition**, that considers who defined the policy. In this sense, we can have policies defined either by the workflow design framework (*predefined policy*) or by the web designer (*user-defined policy*, also known as *compensation chain*).
- **policy execution**, that considers whether the policy is applied in an automated way (*automatic policy*) or in a manual way (*manual policy*). In the former case the policy is automatically applied by the workflow engine after an exception is caught and the engine detects a halted activity. In the latter case a user (the activity executor or another suitable user) can choose the policy to execute through a Web interface (*recovery page*), which is eventually reached after the activity interruption, through an explicit login of the user (Fig. 10).

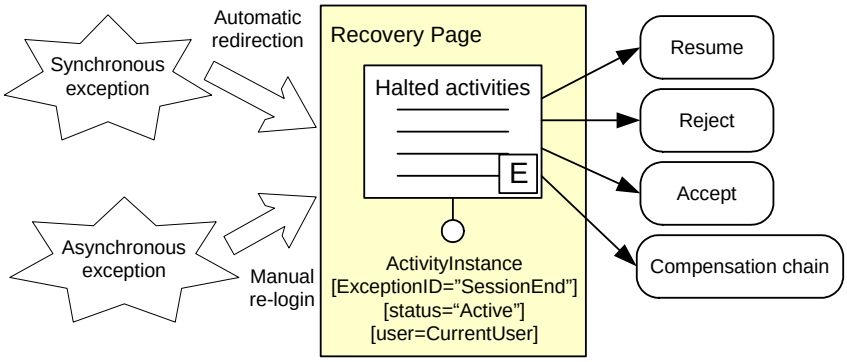

Fig. 10. Manual policies for synchronous and asynchronous exception management

6.2 Policies for Synchronous and Asynchronous Exceptions.

Policy application can be affected by the type of the exception to be managed. In particular, we will apply different policies depending on the fact that exceptions are synchronous or not.

When a synchronous exception occurs the user session is still active. To take advantage of this fact we consider only manual policy for synchronous events: when the exception occurs the user will be redirected to a recovery page and will choose the most appropriate policy (either predefined or user-defined) for the halted activity.

When an asynchronous exception (i.e., a Session End) occurs, the user session is not connected any more and it is not possible to immediately apply manual policies. Therefore we consider both automatic and manual policy for asynchronous events.

Automatic policies are applied automatically and transparently to the user, while manual policies are applied by the user itself, when he starts a new session through a new login. At that point the user

can go in a recovery page and choose the best policy to apply. This behavior is depicted in Fig. 10, together with the predefined policies that are described in next section.

6.3 Predefined policies

Our framework offers three predefined policies: Accept, Reject and Resume. To better understand their behavior, we consider a very simple example, consisting in the order payment activity, as described in Fig. 11: the activity starts, a payment is created and connected to the order, and the user fills up a form with his credit card data. Then, the payment is performed (through a black-box service) and the payment status is updated. If an exception occurs the current step of the activity will be step 1 (since it's the only step of the activity).

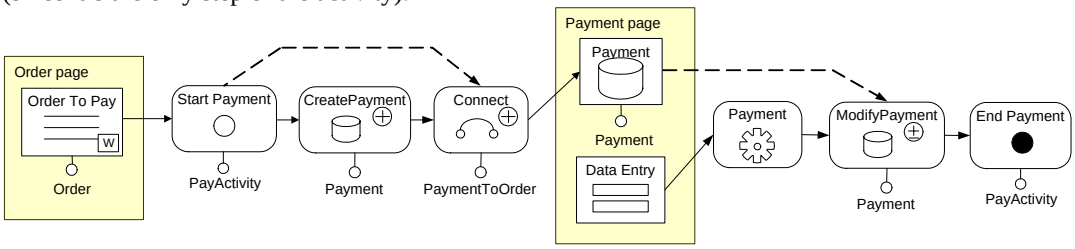

Fig. 11. Payment activity. There is just one step (the payment page), a preceding chain of operation (comprising the create payment unit and the connect unit) and a following chain (comprising the unit for the payment and the modify payment)

For each predefined policy, besides their specific behavior, we assume that they automatically perform the proper updates of the Status of the Exception Instance they are managing, by setting its value into the Exception metadata model to (i) *Solving*, at the beginning of the policy, and to (ii) *Solved* at the end of the policy execution.

Accept policy. It accepts the operations already done by the halted activity, setting the activity status to Completed, executing all the data assignment and activating all the proper following activity. The process can proceed, but it may happen that part of the halted activity was not executed. The accept policy is a forward policy, since it tries to bring the workflow application to a correct status not previously visited, by simply assigning the status Completed to the halted activity.

This policy is suitable only for activities that have some non critical parts, which can be omitted. In all the other cases, it has resulted ineffective, since it leaves the activity results meaningless, thus damaging the whole process case execution. For example, suppose that an exception occurs in the payment activity in Fig. 11. Current step is 1, and if we apply an accept policy, we will consider the activity executed even if the payment unit has not been performed. The process will be enabled to continue, even if the payment has not been performed. Therefore, in this case the accept policy is not a correct choice.

Reject policy. It deletes the data created by the activity, trying to recover the initial state of the database before the activity execution, and assigns the Inactive status to the activity. The reject policy is a backward policy, since it removes the data created by the activity (and all the relationships with connected objects), tries to recover the initial state of the database before the activity execution, and assigns the Inactive status to the activity. Reject policy is not a full rollback mechanism, since not all the operations executed by the activity are undone (i.e., deletion and modification results are kept as they are). Indeed, we don't want to implement a transactional system, with data versioning and so on.

In this way, this policy can be implemented simply by means of a "Created" relationship that connects the Activity Instance to the objects created in the activity itself (an example can be seen in Fig. 4). Once the reject is invoked, the activity is set to Inactive (ready to start) and all the Created objects are removed. Thus, reject is an approximate recovery of the initial state of the activity. This behavior partially limits the effectiveness of the policy but improves its efficiency, avoiding a performance burden resulting from a complete track of all the operations of the activity. Reject policy is suitable for all the activities that should be completely performed, and whose core task consists in creation of objects. With this policy, users can be asked to complete the activity repeatedly, until it is successfully finished. From empiric evaluations, this case results to be very frequent.

If we go back to the payment activity example (Fig. 11), by applying the reject policy in case of exception, we will delete the created instance of payment entity and the instance of the relationship PaymentToOrder, thus canceling the effect of both the create payment unit and the connect unit. The activity is ready to be restarted and the data are in a consistent status.

Resume policy. This policy lets the user resume browsing from the last visited page of the activity before the failure. This policy can be applied only by manual choice of the user, while the first two can be applied both automatically and manually. Browsing is resumed from the last page of the activity generated by the server, i.e. the current step. Note that operations with side effect are not improperly triggered by this policy: if the side effect occurs between the previous and the current page, it is not executed twice, because the user is provided with the url pointing directly to the page; if the side effect occurs after the current page, it has not been executed yet, otherwise the current page should point to the next one.

Resume is a backward policy, since it brings the application and the workflow to a correct state that was previously visited. Indeed, if an exception occurs, either the user session is expired or the page that is shown to the user is different from the current step. The user cannot proceed with the execution of the correct activity, and the whole process status is incorrect. As we said before, the resume policy can only be applied through manual intervention by the user. This can be achieved by providing to the user a recovery page, in which he can see the activities in incorrect status, and can decide to resume them. By reloading the last page generated by the server on the user browser, the activity execution can proceed (e.g., in Payment activity example in Fig. 11, the resume policy lets the user reload the payment page and complete the payment).

6.4 Compensation chains

To allow a more fine-grained exception handling, we allow the designer to define his own recovery policies (e.g., sending warning emails to users, or implementing full rollback capabilities for specific activities). This solution will be adopted to manage the most critical activities only. The user can define operation chains that are triggered by exceptions. This approach exploits the already presented CatchEvent unit, whose purpose consists in capturing a specific exception for a specific activity type and thus triggering a chain of operations. Notice that this requires the designer to explicitly include the update operations for the status of the ExceptionInstance within the compensation chain. Fig. 12 shows a sketched example of compensation chain for the Payment activity depicted in Fig. 11. If any exception arises (no selector condition is specified on the type of exception), the compensation chain is triggered and the following operations are performed: the Status of the ExceptionInstance is set to "Solving", the payment is canceled, the information about it is updated into the database, the Payment activity is closed and finally the ExceptionInstance is set to "Solved".

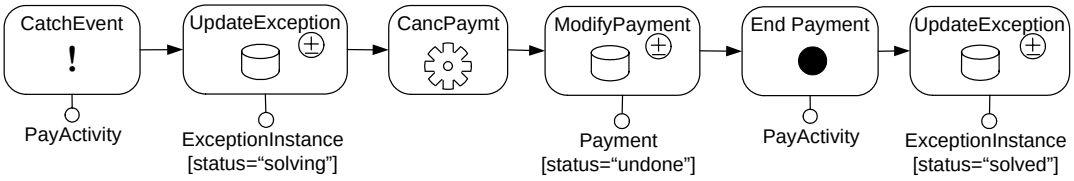

Fig. 12. Payment exception compensation chain. The payment is canceled and the information about it is update into the database. Exception status is updated accordingly.

7. Implementation experience

The concepts presented in this paper have been proved valid on the field, since a prototype implementation has been developed and used to design sample applications. The implementation extends a commercial tool called WebRatio[19], which allows to design and automatically generate Web applications from WebML models. Our extension provides the workflow metadata schema and all the units presented in Section 2.3. Moreover, new units for granting automatic policies enactment are available (Accept, Reject and Resume units).

Several case studies exploiting exception handling capabilities have been implemented, thus validating and refining the approach. The results of this research, which is part of the WebSI project, funded by the EC's 5th framework, has been used by the partners of the WebSI project for pilot applications, and by other projects.

Among them, Acer Business Portal (that includes remote service calls for providing location and driving information to users, and workflow-based interaction between Acer and its commercial partners), and MetalC project[13], which is the most complex among the application we have developed, since it includes a set of B2B portals (one for each business partner). The purpose of the project is to allow business interactions between Italian companies of the mechanical field by means of their respective Web portals, through Web services calls. In this context, complex workflow interactions have been put in place, to grant reliable cooperation. For example, the purchasing process in a B2B scenario consist of a very complex set of interactions, since the buyer typically asks for a quote, the seller makes his offer, then the buyer sends his order for the best offer. In this context, exceptions management becomes very critic. In the implemented communication platform all the discussed recovery policies have been used. Some examples follows: *(i)* if an exception occurs within the AskForQuote activity, an accept policy is performed, and the request is sent even if not all the data are submitted (less relevant data are left in the last steps of the activity); *(ii)* if an exception occurs within the SendOrder activity, the reject policy is applied: data created within the activity is deleted, and the user is asked to restart it; *(iii)* in case of exception within the self-registration activity, which is a long sequence of data submission by the partners, resume policy is exploited, to allow the user resume the self-registration from the point in which he left the application.

An example of user defined recovery becomes necessary within the shipping confirmation activity: once the order has been confirmed and the goods are ready to be shipped, the seller must notify the buyer about the sending. If an exception occurs during the execution of this activity, a user-defined compensation chain is performed, automatically executing the remaining steps of the activity.

8. Conclusions

In this paper we proposed a conceptual approach to exception handling within workflow-based Web applications, described through a metadata model and a set of primitives to be used into hypertext specification. To manage critical situations, we proposed an approach based on exception handling (some Java implementation already exists that could be used to support this approach [21]), and definition of predefined and user-defined policies, that have been tested on the field.

The main advantage of our approach stands in allowing the definition of exception handling and compensation chains without lowering the abstraction level of the design.

Future work will address refinement of the implementation, to allow a more seamless and transparent integration of exception handling within WebML specification, to avoid the need of explicitly specifying in WebML all the basic steps of exception handling. A second research direction is towards study of exception handling in remote service calls. Some preliminary considerations have been done, and we expect an approach similar to the one we have studied for workflow-based Web applications.

References

1. Atzeni, P., Mecca, G., Merialdo, P.: Design and Maintenance of Data-Intensive Web Sites. EDBT 1998: 436-450.
2. Baresi, L., Garzotto, F., Paolini, P.: From Web Sites to Web Applications: New Issues for Conceptual Modeling. ER Workshops 2000: 89-100.
3. Brambilla, M., Ceri, S., Comai, S., Fraternali, P., Manolescu, I.: Specification and design of workflow-driven hypertexts, Journal of Web Engineering, Vol. 1, No.2 (2002).
4. Canchero, C., Gòmez, J.: Advanced Conceptual Modeling of Web Applications: Embedding Operation Interfaces in Navigation Design, 21th International Conference on Conceptual Modeling, El Escorial, Madrid.
5. Casati, F., Ceri, S., Paraboschi, S., Pozzi, G., Specification and implementation of exceptions in workflow management systems. ACM Transactions on Database Systems, Sept. 1999, (Vol. 24, No. 3), pp. 405-451.
6. Ceri, S., Fraternali, P., Bongio, A.: Web Modeling Language (WebML): a modeling language for designing Web sites. WWW9/Computer Networks 33(1-6): 137-157 (2000).
7. Ceri, S., Fraternali, P., Bongio, A., Brambilla, M., Comai, S., Matera, M.: Designing Data-Intensive Web Applications, Morgan-Kaufmann, December 2002.
8. Conallen, J.: Building Web Applications with UML. Addison Wesley (OTS), 2000.
9. Hagen, C., Alonso, G.: Exception Handling in Workflow Management Systems, IEEE Transactions on software engineering, October 2000 (Vol. 26, No. 10), pp. 943-958
10. Koch, N., Kraus, A.: The Expressive Power of UML-based Engineering, Second International Workshop on Web Oriented Software Techonlogy, CYTED 2002, 105-119.
11. IBM MQSeries Workflow Homepage: http://www.ibm.com/software/ts/mqseries/workflow/v332/
12. Oracle Workflow 11i: http://www.oracle.com/appsnet/technology/products/docs/workflow.html
13. MetalC project Homepage: http://www.metalc.it
14. Miller, J. A., Sheth, A. P., Kochut, K. J., Luo Z. W.: Recovery Issues in Web-Based Workflow, CAINE-99, Atlanta, Georgia (November 1999) pp. 101-105.
15. Schwabe, D., Rossi, G.: An Object Oriented Approach to Web Applications Design. TAPOS 4(4): (1998).
16 Troyer, O., Casteleyn, S.: Modeling Complex Processes for Web Applications using WSDM, Third International Workshop on Web Oriented Software Technology, Oviedo 2003, 1-12.
17. Rossi, L., G., Schmid, H., Lyardet, F.: Engineering Business Processes in Web Applications: Modeling and Navigation Issues, Third International Workshop on Web Oriented Software Technology, Oviedo 2003, 81-89.
18. WebML Project Homepage: http://www.webml.org
19. WebRatio Homepage: http://www.webratio.com/
20. Workflow Management Coalition Homepage: http://www.wfmc.org
21. Ofbiz WF Java implementation: http://www.ofbiz.org/api/components/workflow/build/javadocs/