An analysis of modeling flaws in HL7 and JAHIS

Eduardo Fernandez
Department of Computer Science and Engineering
Florida Atlantic University
Boca Raton, FL 33431, USA
ed@cse.fau.edu

Tami Sorgente
Department of Computer Science and Engineering
Florida Atlantic University
Boca Raton, FL 33431, USA
tsorgente@fau.edu

ABSTRACT
Health Level 7 (HL7), one of the ANSI standards organizations has a mission to provide standards for the execution, management, and integration of data to support clinical patient care. HL7 will play an important role in the implementation of the HIPAA regulations. HL7 has developed a Reference Information Model (RIM), an object-oriented model of clinical data. JAHIS is a Japanese organization that has developed extensions to this RIM. Instead of using the Unified Modeling Language (UML), the standard notation for object-oriented software development, these two organizations have developed specialized object-oriented models. This has resulted in languages which are incompatible with the current use of UML. The consequences of this choice are the loss of the possible use of a large variety of existing models and patterns. What is worse, it will be difficult to add security specifications in their models, a critical aspect in the electronic interchange of medical records. We discuss here the shortcomings of HL7 and JAHIS as modeling languages and as languages in which to add security specifications. We also propose some solutions to this situation.

Categories and subject descriptors: D 2.2 Object-oriented design methods, D 3.3 Language constructs and features

General Terms: Design, Security

Keywords: Healthcare information systems, HL7, medical systems security, object-oriented design, software patterns, security, UML

1. INTRODUCTION
Health Level Seven (HL7) is one of several ANSI-accredited Standards Developing Organizations (SDOs). The mission of HL7 is to provide standards for the exchange, management and integration of data that support clinical patient care, and the management, delivery and evaluation of healthcare services [8]. HL7 has developed a Reference Information Model (RIM), which is a graphical representation of clinical data (domains). They use an object-oriented technique to develop the models. This model identifies the life cycle of the events that messages will carry. HL7 has been adopted as a standard and it will play a basic role in the enactment of HIPAA regulations in the US and in similar laws in other countries.

Four Japanese organizations [16] including JAHIS (Japanese Association of Healthcare Information Systems Industry) have developed Reference Enterprise Models (REM) using RM-ODP (Open Distributed Processing - Reference Model) and UML (Unified Modeling Language) for Japanese hospital information and Japanese healthcare domains. These are very detailed healthcare information models (HIS). The HIS are used to “divide and conquer” the complex issues and describe objects, roles and communities (departments). This project developed a model based on HL7 for the Radiology department.

The purpose of this paper is to show that the HL7 modeling approach and the JAHIS additions don’t conform to the Unified Modeling Language (UML) semantics and have serious flaws from a conceptual modeling perspective. This situation will bring a variety of problems in their future use, including: A large variety and number of existing models, patterns, and programs cannot be directly applied without cumbersome transformations. Security approaches at the application level, expressed as patterns or similar models will not be applicable, resulting in the use of low-level security mechanisms, an approach that has repeatedly failed to produce secure systems.

We elaborate these two points in this paper, concluding that HL7 and JAHIS are not appropriate languages for the needs of current and future medical systems. Section 2 presents some background, while Section 3 brings up some simple UML models and shows that they are difficult to describe in HL7 concepts. Section 4 shows that while it is easy to apply security patterns to UML models, it is not easy to do similarly in HL7 or JAHIS models. Section 5 evaluates these models based on the previous discussions. We end with some conclusions and a sketch of a possible solution to the problems posed by HL7.

2. BACKGROUND

2.1 Object-oriented design and patterns

Two important advances are having a significant effect on the quality and security of software applications:

• The increasing use of object-oriented design, applied through the UML, which has expanded the range of size and complexity for software systems [14].
• The emergence of patterns that summarize the experience and knowledge of designers. There are now catalogs of analysis patterns that describe conceptual models, catalogs of architectural patterns that describe system architecture aspects, and catalogs of design patterns that show how to make software more flexible and adaptable [12].

Any modern software system of high complexity must be able to use these advances or it will fall in the position where it needs to reinvent the wheel and will not be able to produce high quality software. This fear probably lead the designers of HL7 and JAHIS to use a UML-like type of modeling. However, as we will see, they fell short in that they introduced an idiosyncratic approach that is not compatible with standard use of UML. We describe the salient features of HL7 and JAHIS in order to show how they differ from standard UML.

### 2.2 Basic RIM classes

HL7 uses a message definition methodology, HL7 V3 standard development process, to develop a network of inter-related models that depict the static and behavioral aspects of the design and for requirements of its standards. These standards are for the exchange, management, and integration of the data that support’s clinical patient care and the management and evaluation of healthcare services. The HL7 RIM is the root of all information models and structures developed as part of the V3 development process [8]. There are six “back-bone” classes of the RIM model, detailed in Table 1.

<table>
<thead>
<tr>
<th>Class name</th>
<th>definition</th>
<th>example of HL7 RIM</th>
<th>Sub-classes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Entity</td>
<td>Represents the objects that take part in healthcare</td>
<td>person, medical facility, facility asset</td>
<td>yes</td>
</tr>
<tr>
<td>Role</td>
<td>Demonstrates the competency of the Entity as it participates in a medical Act</td>
<td>doctor, patient, nurse</td>
<td>yes</td>
</tr>
<tr>
<td>Act</td>
<td>Represents the actions or acts that are performed or executed. An Act is performed by an Entity acting as a Role</td>
<td>clinical observation, check location availability, and discharge a patient.</td>
<td>yes</td>
</tr>
<tr>
<td>Participation</td>
<td>Expresses who performed the act, for whom the act was done, etc. It is an association between an Act and a Role with an Entity playing that Role</td>
<td>Entity: DoctorA, Role: Doctor, Act: discharge a patient, Participation: participation of a doctor</td>
<td>no</td>
</tr>
<tr>
<td>ActRelationship</td>
<td>Represents the binding of one act to another, such as the relationship between an order for an event and the occurrence of the event</td>
<td>relationship between an order for an event and the occurrence of the event</td>
<td>no</td>
</tr>
<tr>
<td>RoleLink</td>
<td>Represents relationships between individual roles</td>
<td>links the Physician’s relationship with an organization and a patient’s relationship with the organization to express the patient /physician relationship</td>
<td>no</td>
</tr>
</tbody>
</table>

The classes are linked either by a set of association relationships, identified by unique role names, or by generalization relationships as shown in Figure 1 [8].

![Figure 1. Major classes of HL7 RIM](image-url)
The **Entity, Role, Participation, Act, ActRelationship**, and **RoleLink** classes of the RIM model are used together as a pattern. Acts connect to Entities in their Roles through Participations and connect to other Acts through ActRelationships. Acts have several subclasses. The Entity hierarchy encompasses living subjects (including human beings), organizations, material, and places and their specializations. It does not indicate the roles played, or acts that these entities participate in (e.g., patient, provider) [8]. An Entity participates in an Act in a particular Role. Most attributes of Role are attributes of the playing entity while in the particular role.

Acts are the pivot of the RIM; all domain information and processes are represented primarily in Acts. ActRelationships and RoleLinks are used to describe relationships. In HL7 terms, a medical record is a documentation of the individual actions that make up the diagnosis, treatment and care of a patient. The Act “**mood**” code specifies whether the Act is an activity that has happened, can happen, is happening, is intended to happen, or is requested/demanded to happen [2].

### 2.3 HL7 RIM Representations

The foundation of HL7 modeling is defined in a UML profile, however they use a graphic representation that better represents their “sentences.” They use square boxes to represent nouns, displaying all attributes and constraints. The verbs are shown as arrows linking the nouns. Consider the following scenario: A clinician orders a lab test for one of her patients. The test will be performed on a specimen collected at her office. She will send the specimen by courier, and expects to receive a confirmation that the test will be performed, and receive a result of the test [2]. The HL7 diagrammatic representation is as displayed in Figure 2 uses HL7 classes as objects.

In Figure 2 ObservationOrder and Order are **Acts.** Author, Performer, Subject and Record target are **Participations.** Physician, Laboratory, Specimen, and Patient are **Roles.** Definition, located between Order and Ordered test is an **ActRelationship.**

![Figure 2. HL7 Laboratory Observation Order (from [2])](image-url)

---

### 2.4 JAHIS RM-ODP enterprise language and UML

The REM was developed as a joint team project. The team was comprised of experts from JAHIS, INTAP (Interoperability Technology Association for Information Processing, Japan), and MEDIS (The Medical Information System Development Center). Together, they used the RM-ODP (Open Distributed Processing - Reference Model), a family of standards from ISO/IEC JTC1 and ITU-T. They found this technique effective in modeling complicated systems and clearly defining the requirements. However, according to JAHIS, RM-ODP defines what should be described in each viewpoint, the resulting specifications are not sufficient from a software engineering point of view. To compensate for this limitation, the modeling language UML was introduced and used in combination [11].

The functions of the ODP system are intentionally not analyzed in detail: the focus is on the Enterprise Viewpoint (objective, scope, and policy) specification. There are four scenarios modeled which include:

- Initial visit (outpatient)
- Ambulatory care (outpatient)
- Hospital change (discharge)
- Hospital admission (inpatient)

For the REM the enterprise objects include administrators, physicians, staff, family, etc. The roles of the project include outpatient, administrator, visitor reception staff, and others. For the dynamic aspects of this project and the description of scenarios, JAHIS used a form of activity diagram as displayed in Figure 3 that describes outpatient admission.

![Figure 3 JAHIS activity diagram for outpatient admission (from [11])](image-url)

---

### 3. COMPARING MODELING CONSTRUCTS

#### 3.1 Patient records

To compare approaches we look at the patient record. Figure 4 describes some concepts of the Patient record in the form of objects. As indicated earlier, the foundation of HL7 modeling is defined in a UML profile but they use a graphic representation that better represents their “sentences” [2].

In Figure 4 the objects are located around the **ActContext AC_Clinical_document.** The Role **R_Patient** participates as a **P_Patient** with **AC_Clinical_document**. Role of **R_Patient** is...
associated with the Entity **Person_as_Patient** and also associated with the Entity **Organization**, representing the facility for treatment. Figure 4 details the Patient relationship with the clinical document.

In the Patient Record Management Pattern of Figure 5 we use UML to represent similar information. The **Patient** class aggregates **Record** and **Record** aggregates **Treatment_Instance**.

It is easy to see that the UML representation is much clearer than the HL7 description. Constraints on the classes and their variables can be described using OCL as shown later.

![Figure 4. HL7 Patient and Clinical Document Relationship [1]](image)

![Figure 5. Patient Record Management Pattern](image)

### 3.2 JAHIS Reference Information

**Model and Patient Treatment Patterns**

JAHIS developed an extensive study of medical information modeling and arrived at a complex information model. According to their report, the model depicts the entire structure of the radiology department logically and statically Figure 6 represents a section of the JAHIS project information model describing the radiology department. For comparison purposes we concentrate here on the patient information area of the model.

![Figure 6. JAHIS Information model section (from [10])](image)

219
There is extensive use of aggregation in this model. One example is having both the Physician information and the Patient information described as aggregates of the Exam order. Their aggregations do not correspond to a conceptual view. **Physician information** is not a conceptual component of an **Exam order**. As another example, **Patient information** aggregates **Room information**. In comparison, UML describes the relationship between the room or location assigned to the patient as shown in Figure 7. In this model aggregations are used to describe semantic aspects of the model, not representational aspects as in JAHIS.

### 3.3 Dynamic Representations

Part of the JAHIS project was the development of “object models.” They used their version of UML to develop these models both from their information model and HL7 RIM. First, we analyze some aspects in the development of inpatient admission and the assignment of a location to better understand HL7. The Act of assigning a location to an inpatient is a **PatientEncounter**. The following information is from HL7 documentation.

**Class: PatientEncounter (in Acts)**

PatientEncounter is a specialization of: Act

**Definition of PatientEncounter:** An interaction between a patient and care provider(s) for the purpose of providing healthcare-related service(s). Healthcare services include health assessment. **Examples:** outpatient visit to multiple departments, home health support (including physical therapy), *inpatient hospital stay*, emergency room visit, field visit (e.g., traffic accident), office visit, occupational therapy, telephone call.

**Entities:**

Definition of Organization: An Entity representing a formalized group of entities with a common purpose (e.g. administrative, legal, political) and the infrastructure to carry out that purpose. **Examples:** Companies and institutions, a government department, an incorporated body that is responsible for administering a facility, an insurance company.

Definition of Person: A subtype of LivingSubject representing a human being.

**Constraints:** This class can be used to represent either a single individual or a group of individuals based on the value of Entity.determinerCode and Entity.quantity.

**Roles:**

Definition of Employee: A role played by a person who is associated with an organization (the employer, scoper) to receive wages or salary.

**Discussion:** The purpose of the role is to identify the type of relationship the employee has to the employer rather than the nature of the work actually performed (contrast with AssignedEntity).

**Definition of Patient:** A Role of a LivingSubject (player) as a recipient of health care services from a healthcare provider (scoper) [8].

The Act classes describe a very microscopic view of actions. The use cases are not defined. The Person, a subclass of Entity, constraints are represented as codes. Next, we define an HL7 relationship using objects for each major class as in Figure 8.
A subsequent report for the JAHIS radiology department was developed in combination with the HL7 message exchange RIM model. By using HL7 RIM, every object could be depicted by filling a value for each attribute based on the four abstracted classes (Entity, Role, Participation, and Act) as we did in Figure 8. A series of complex “Object Diagrams” like the one in Figure 9 were produced. This diagram describes the Entities, Roles, Participations, and Acts involved in ordering an x-ray examination. We have enlarged for detail a representation of the relationship between the objects defined as:

Entities: Hospital, Physician
Roles: Hospital, Physician: Qualified practitioner
Act: Hospital reception: Patient_encounter
Participation of Physician
RoleLink: Hospital-Physician

![Figure 8 HL7 Relationship for assigning a location](image)

![Figure 9. JAHIS version of HL7 object model for scenario of placing an order for a general x-ray examination](image)

4. MODELING SECURITY
Defining security constraints in medical patient systems is complex. There are a variety of policies that need to be represented, such as: patients can see their own records, doctors are allowed to modify the records of their patients, patients must be notified of the use of their records, etc. Representing policies like these in HL7 RIM would be extremely cumbersome. Using UML, the representation of these policies is rather straightforward [6]. Another advantage of a model of this type is that to enforce such security restrictions one needs the participation of the lower levels. For example, these classes may correspond to XML documents. In that case, we can apply XML security standards and patterns to build secure systems. This is again very difficult to do in HL7 RIM style models.

As an illustration, Figure 10 shows a model that describes policies such as the ones mentioned above. It also shows records related to each other as for tracking infectious diseases, doctors as custodians of some records, and patients authorizing the use of their records. These policies can be made more precise by using OCL.
5. EVALUATION OF THE MODELS

We summarize now some conclusions based on the examples shown and our study of the HL7 and JAHIS models.

HL7 models

HL7 is complex and designed for implementation, not for conceptual modeling. We show below some of the sources of complexity. Although the designers deny having a software orientation, it is clear that the model contains artifacts normally used in the design stage of software systems. A measure of their complexity is that their documentation is hard to understand.

- **Ad hoc variation of UML**—Making it incompatible with the conceptual models already developed for medical systems. There are a variety of patterns that cannot be used directly, including security patterns.

- **Roles are job descriptions**—Roles indicate competency, e.g., a doctor, a nurse. This is not the typical use of role in security. In the Role-Based Access Control (RBAC) model, the most common in current systems, roles are tasks with a set of assigned rights. A doctor has only one competency but can have several roles of the latter type, e.g., admitting doctor, primary doctor, consulting doctor, etc. A role of the latter type is realized in HL7 through its class Participation.

- **Entities**—Having and explicit concept of entity is not appropriate for conceptual modeling. In a conceptual model everything is an entity, such a class doesn’t add anything to the semantics of the model. Here it is being used in the sense of ‘object’ in Java and similar languages, a class that contains common aspects (usually implementation) of all classes. This is confirmed by seeing that the attributes in Entity are software related, e.g., codes and multiplicities. A conceptual model that lumps people and physical objects together is clearly not very semantically precise.

- **Associations have no names and no semantic value**—Associations in HL7 have no semantic value. Their semantics is placed in a separate class, a RoleLink or an ActRelationship. UML associations have semantic value and association classes can be used to add details to links.

- **States are hidden as mood codes**—UML uses state diagrams to describe the stages in the lifecycle of an object. In HL7 these stages are implicitly encoded using mood codes. This precludes any analysis of state correctness and correlation with sequence (scenario) diagrams.

- **Cumbersome and inefficient**—As a result of its misuse of associations their models become cumbersome and unnecessarily complex, as it can be seen in Figures 2, 6, and 9. Models that in UML take a few classes and associations require a much larger number of classes in HL7.

- **Software-dictated names**—Because the need to separate the different types of classes in their model and these classes can take up arbitrary names, they separate them using a prefix letter, as in P_patient, to indicate a class of type Participation. This can be done much more naturally in UML by using stereotypes, e.g., <<Participation>> Patient. Stereotypes allow easy extension, one can define arbitrary stereotypes, not being subject to a predefined set.

- **Loss of the possibility of using security patterns and models**—Given the importance of security when medical records are fully computerized and exchanged through the Internet, this may be its most serious flaw.

Some of these deficiencies were already noted by Mead [13], who said that the following may be true about HL7:

- **States are hidden as mood codes**—UML uses state diagrams to describe the stages in the lifecycle of an object. In HL7 these stages are implicitly encoded using mood codes. This precludes any analysis of state correctness and correlation with sequence (scenario) diagrams.

- **Cumbersome and inefficient**—As a result of its misuse of associations their models become cumbersome and unnecessarily complex, as it can be seen in Figures 2, 6, and 9. Models that in UML take a few classes and associations require a much larger number of classes in HL7.

- **Software-dictated names**—Because the need to separate the different types of classes in their model and these classes can take up arbitrary names, they separate them using a prefix letter, as in P_patient, to indicate a class of type Participation. This can be done much more naturally in UML by using stereotypes, e.g., <<Participation>> Patient. Stereotypes allow easy extension, one can define arbitrary stereotypes, not being subject to a predefined set.

- **Loss of the possibility of using security patterns and models**—Given the importance of security when medical records are fully computerized and exchanged through the Internet, this may be its most serious flaw.

JAHIS

Some of the same errors propagated to JAHIS. They corrected some errors and their approach is more UML-like, but most of the same problems remain.
• **Predefined roles (jobs)**—This is good in some sense, it facilitates defining some models. On the other hand, more roles may be needed.

• **Lack of conceptual clarity and too much complexity**—Unnecessary complexity remains. There is no clear semantic definition of their models.

• **Development of the relationships between the objects is complex and the Reference information models use hierarchies of aggregation**—As seen in Figure 6, JAHIS models mostly use aggregations instead of regular associations. These aggregations are really software constructs, they do not correspond to semantic associations. For example, the Room Information is not a component of the Patient Information, it should be related through a standard association. All this results in a much larger number of classes than necessary.

• **Object diagrams**—These are used to describe scenarios. The standard construct for scenarios in UML are sequence or collaboration diagrams, which show clearly the sequence of interactions between objects. HL7 scenarios are similar to collaboration diagrams but without their well-defined semantics.

• **Confusion of dynamic and static models**—Similar to HL7.

• **Diagrams closer to UML in appearance but still with different semantics**—Their object models use UML in a very ad hoc way as described in the items above.

6. CONCLUSIONS

UML models and analysis patterns describe the information in a very clear conceptual way. With these models we are able to use methodologies to combine patterns and add security. Both HL7 and JAHIS diagrams become so complex that modeling and adding security is difficult and unclear. In contrast to UML, neither HL7 nor JAHIS has developed formal or semi-formal semantics for their models. Short of replacing HL7 and derived models with a full-UML model, we see another possibility: build a catalog of HL7 patterns with their corresponding UML equivalents. We indicated earlier that combinations Entity, Role, Act, ActRelationship, and Participation are used as a pattern. We can find a UML equivalent for this type of pattern and apply it to specific cases.

REFERENCES


