

An Electrocardiogram (ECG) Domain Ontology

Bernardo Gonçalves¹, Giancarlo Guizzardi^{1,2}, José Gonçalves Pereira Filho¹

¹Computer Science Department, Federal University of Espírito Santo (UFES),
Vitória, Brazil

² Laboratory for Applied Ontology (ISTC-CNR)
Trento, Italy

{bgoncalves,zegonc}@inf.ufes.br, guizzardi@loa-cnr.it

Abstract. *This paper presents an ontology of the electrocardiogram domain. The ontology purpose is to bring in a theory of the electrocardiogram (ECG). The ECG is the most applied test for mapping the heart activity. Notably with the recent advances in information and communication technologies, new services in Healthcare have been provided, standing out remote reporting and telemonitoring. In this context the ECG has a key usefulness. However, the diversity of biomedical data combined to heterogeneous information systems have been inhibited these advances in Healthcare environments. This initiative then addresses the need for semantic interoperability, in general, and data integration, in particular, in Health Informatics. Besides, it also provides an approach for developing ECG automatic analysis systems.*

1. Introduction

In 1902, Electrocardiography was introduced by Einthoven following his invention of the string galvanometer, for which he received the Nobel Prize. Electrocardiography is the technique of recording the electrical signal generated the heart activity. This record is called electrocardiogram (ECG) [Geselowitz 1989], [Guyton & Hall 1996]. The ECG is the most frequently applied test for measuring the heart activity in Cardiology. Compared with other examination procedures, it is fast, cheap and non-invasive. The importance of the Electrocardiography is remarkable since heart diseases constitute one of the major causes of mortality in the world. According to estimates, more than 100 million ECGs are recorded annually in the Western Europe [SCP 2002].

Since Einthoven's invention, the measurement of bioelectric potentials generated in the human heart has been object of research by biomedical engineers leading to many improvements in instrumentation and digital computers. According to [Geselowitz 1989], ECG was possibly the first diagnostic signal to be studied with the purpose of automatic interpretation by computer programs. There are two characteristics for interpretation in the ECG: (i) the morphology of waves and complexes which compose a cardiac cycle; and (ii) the timing of events and variations in patterns over many beats. In this way, the analysis of the ECG waveform supports identifying a wide range of heart diseases. The characterization of each cardiopathy manifests itself by specific modifications on the characteristics (i) and (ii).

In recent years, by taking advantage of information and communication technologies (ICT), the Telecardiology has found in the transmission of electrocardiogram (ECG) an

economic and efficient way of rendering remote medical services. These services range from simple remote reporting through the transmission of a simple ECG record (by means of the latest mobile and wireless technologies) and the transmission of the ambulatory ECG (AECG) to the possibility of providing heart telemonitoring anytime, anywhere [Salvador et al. 2005], [Andreão et al. 2006].

In face of all this, the storage and transmission of ECG records have been object of some initiatives concerning standardization. Among the reference standards, AHA/MIT-BIH [Goldberger et al. 2000] and SCP-ECG [SCP 2002] were conceived regarding record's storage and transmission respectively. Furthermore, on account of the Internet popularization new standards have been conceived in order to integrate interoperable and user-driven solutions, standing out HL7 [HL7 2003], FDADF [Brown et al. 2002] and ecgML [Wang et al. 2003]. Nevertheless, despite the fact that such ECG formats are considered reference standards in Cardiology, these standards are not amenable to foster semantic interoperability, in general, and data integration, in particular, among ECG-based applications. This is either because they are strongly-coupled with specific programming languages and environments, not human readable (an important requirement to permit the analysis of the data from electrocardiography's domain experts), their metamodels mix-up domain concepts with presentation ones, or because they are represented as conceptual models of poor expressivity and clarity.

With such issues in mind, we have developed an ontology of the ECG domain. The ontology purpose is to provide a theory of the ECG in order to: (i) address semantic interoperability and data/standard integration between health systems, which is the currently key point of the research on Health Informatics [Ackerman et al. 2002]; (ii) allow a novel artificial intelligence approach for ECG automatic analysis; and (iii) convey a knowledge repository about ECG. The proposed ontology covers the domain from multiple complementary viewpoints. In this way we provide an integral view of the heart activity from the microscopic perspective to the macroscopic one.

The ECG ontology has been developed in the context of the TeleCardio project [Andreão et al. 2006]. TeleCardio is a telehomecare system for remote monitoring of patients with cardiological syndromes. This system is built on top of a middleware for mobile and context-aware applications as well as mechanisms for ECG signal analysis and for automatic alert generation.

The remainder of this paper is organized as follows: Section 2 provides a brief description of the studied domain; Section 3 introduces the ontological engineering approach employed in this article; Section 4 presents the ECG ontology and its sub-ontologies; Section 5 discusses related work; and finally, Section 6 concludes the paper.

2. A Brief Description of the Electrocardiogram Domain

Bioelectric sources spontaneously arise in the heart at the cellular level. For a short understanding, the heart cells are immersed in a fluid matrix that is separated of the interior of the cells by their membranes. These membranes carry out a control of ions transport. In the resting state, the interior of the cells has a negative potential w.r.t. the exterior, i.e., the cells are electrically polarized. However, particularly in the sinoatrial (SA) node and atrioventricular (AV) node, which are specific regions of the heart muscle (myocardial), cells abruptly depolarize and then return to its resting value. This

phenomenon is a result of ions passing in either direction across the cells' membrane. Therefore, notably SA and AV node cells give rise to an electrical impulse which is propagated to its neighboring cells and normally reach the entire heart. That is why SA and AV nodes are called the heart pacemakers. However, because such electrical impulse arises in SA node at a faster rate and with a higher intensity than in AV node, the impulse generated in the latter is then overdriven [Geselowitz 1989].

For conveying the cardiac electrical impulse generated in the SA node around the heart, there are fiber cells beyond the SA and AV that compose a conduction system (see Figure 1-(i)). The major conduction pathway is called His-Purkinje system, which is composed by the bundle of His (separating the left bundle branch from the right one) and by the Purkinje fibers (indicated as the conduction pathways). Although those fiber cells also depolarize and repolarize, the electrical current generated by them has not a considerable outcome except to convey the cardiac impulse around heart tissues. The return path of this impulse often involves the entire body [Guyton & Hall 1996].

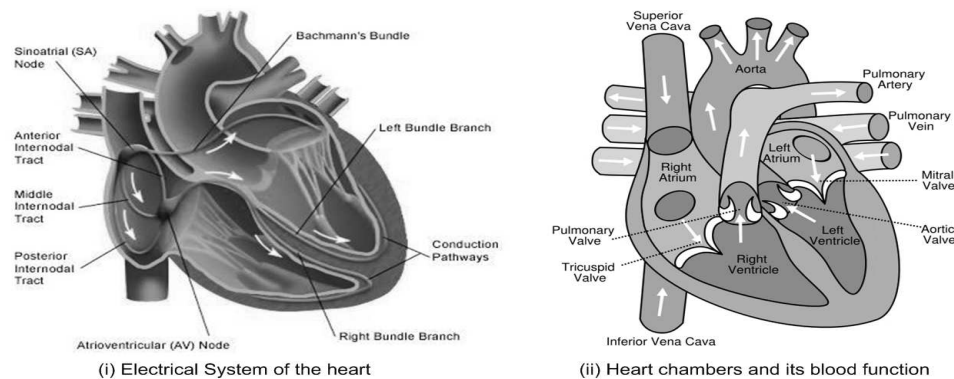


Figure 1. The human heart

As a response for the electrical current that is conveyed around heart tissues, the myocardial holds contractions in atrial and ventricular areas that push blood respectively into the ventricles and either systemic or pulmonary circulation. In view of this blood transportation function by means of muscle contractions, the heart plays the role of a pump. With respect to blood storage, the heart plays the role of a blood container. It is able to carry out this function because it is a chamber, which is divided into four sub-chambers: left atrium, right atrium, left ventricle and right ventricle. The heart key function indeed is pumping blood either to systemic and pulmonary circulation.

On the “left heart”, as the arrows in Figure 1-(ii) indicate, the left atrium receives oxygenated blood from the lungs ready to be moved to the peripherals (head, torso and limbs) for conducting several resources (e.g. oxygen, proteins, etc.) to the body cells. When the left atrium is almost full of oxygenated blood, the mitral valve is opened by the blood pressure and the blood then is pushed to the left ventricle by an atrial contraction¹. The left ventricle in turn moves the blood to the peripherals through a strong muscle contraction. On the “right heart” (Figure 1-(ii)), in the meantime, the right atrium receives de-oxygenated blood from the peripherals to be moved to the lungs for

¹Indeed, before the atrial contraction about 75% of the blood is already in ventricular chamber, remaining only about 25% of the blood which is pushed by the atrial contraction. For this reason, the atrial contraction is not vital for a well heart functioning.

being oxygenated again. When the right atrium is almost full of de-oxygenated blood, the tricuspid valve is opened by the blood pressure and the blood then is pushed to the right ventricle. It moves then the blood to the lungs by a ventricular contraction.

The heart activity can be mapped into an ECG. For this purpose, in a recording session one recording device performs observations evenly spaced in time for measuring electrical potential differences around the patient's body surface. These observations are made at the same time at different electrode positions (leads) through different device channels. These correlated observation series carried out in a recording session compose an ECG record, which provides direct evidence of cardiac rhythm and conduction, and indirect evidence of certain aspects of myocardial anatomy, blood supply and function.

An ECG is composed by one or more cardiac cycles, i.e., heart beats (see Figure 2-(i), source: [ELC 2007]). A cycle, as introduced by Einthoven, has elementary forms, or waves, which he names PQRST and are outlined as P wave, QRS complex and T wave. The P wave and the QRS complex map the electrical potential generated by atrial and ventricular depolarization respectively. Atrial and ventricular muscle contractions start at the peak of these waves. The T wave maps the ventricular repolarization. The atrial repolarization can not be seen in the ECG waveform since its resulting potentials are small in amplitude and are overridden by the QRS complex. There is also a U wave, but its origin is still not completely known. The RR interval is used to measure the duration of a cycle. Through several cycles it is possible to obtain an average of the heart rate w.r.t. the time interval related to these cycles [Geselowitz 1989], [Guyton & Hall 1996].

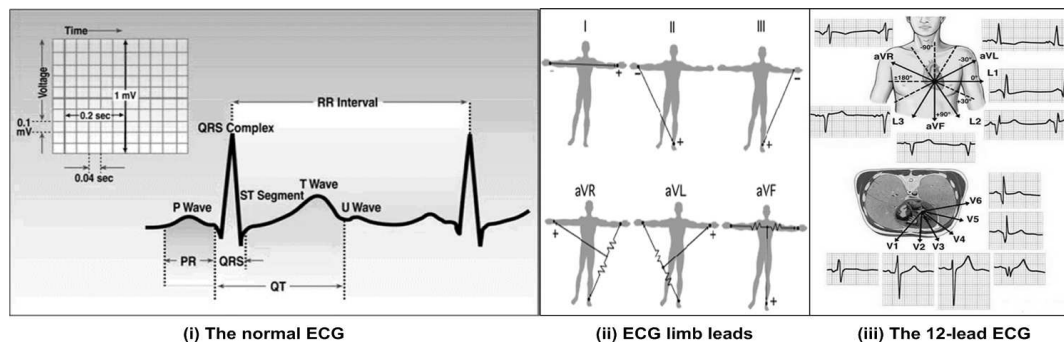


Figure 2. Electrocardiography concepts

ECG data are acquired from different manners for different purposes. The most commonly used method is the so-called 12-lead ECG. In this version, twelve leads are used to acquire the ECG. A lead is a specific arrangement of electrode(s), possibly using weighting resistors, that provides a viewpoint of the heart activity. In combination, multiple leads provide an accurate picture of the heart behavior. The twelve leads standardized in Electrocardiography are divided in limb leads and precordial leads. They are obtained by electrodes attachments as follows. Leads I, II and III are obtained by the potential difference between (i) left arm and right arm, (ii) left leg and right arm, and (iii) left leg and left arm respectively. By convention other leads are obtained through weighting resistors in combination with a common reference electrode placement at the chest called the Wilson central terminal (WCT). They are the augmented limb leads AVL, AVR and AVF; and the precordial leads V1, V2, V3, V4, V5 and V6. The former are obtained by opening the resistor attached to the limb in question. The latter are

obtained by six electrodes placements at heart specified anatomically w.r.t. the rib cage in a suitable way to measure heart electrical signals at multiple angles (see Figure 2-(ii,iii)) [Geselowitz 1989]. For a more detailed description see [Guyton & Hall 1996]. On next section we introduce the approach used for developing the ECG ontology.

3. The Ontological Engineering Approach

In order to develop the ontologies presented in Section 4, we have adopted a combination of two methodologies in the literature. Firstly, we have employed the SABIO method proposed in [Falbo 2004]. This method has been tested for the last ten years in the development of a number of domain ontologies in areas ranging from Harbor Management to Software Process to Media on Demand Management. SABIO prescribes an iterative process comprising the following activities: (i) the identification of the ontology purpose by means of competence questions; (ii) the capture of concepts existing in the domain as well as its relations and its properties; (iii) the ontology formalization, which is the definition of formal axioms by using First-Order Logic (FOL); (iv) the search for existing ontologies with reuse and integration in mind; (v) the evaluation of the ontology for identifying inconsistency as well as verifying the truthfulness with the ontology purpose; (vi) ontology documentation. It is important to highlight that competence questions play a prominent role in this methodology by: (1) defining the scope and purpose of the domain conceptualization being developed; (2) serving as a testbed for ontology evaluation – the competence questions are the questions the ontology are supposed to answer [Falbo 2004].

As discussed in depth in [Guizzardi 2007], as any engineering process, ontology engineering should include phases of conceptual modeling, design and codification. These phases, in turn, shall produce different artifacts with different objectives and, as consequence, require different types of modeling languages and methods with specific characteristics. As argued in that article, in a conceptual modeling phase, an ontology should strive for expressivity, clarity and truthfulness in representing the subject domain at hand. The same conceptual model can then give rise to different ontology codifications in different languages (e.g., F-Logic, OWL DL, RDF, ORM, Ontolingua) in order to satisfy different non-functional requirements such as computational tractability, decidability, etc. In [Guizzardi 2006], the author demonstrates that semantic web languages such as OWL and RDF are codification languages and fall short in serving as a language for the phase of ontology conceptual modeling.

For the reasons just mentioned, in this paper, in order to represent our ECG ontology we have used an ontologically well-founded UML modeling profile proposed in [Guizzardi 2005]. This profile comprises a number of stereotyped classes and relations implanting a metamodel that reflect the structure and axiomatization of a foundational (and, thus, domain independent) ontology named **UFO (Unified Foundation Ontology)** [Guizzardi & Wagner 2005]. A complete description of UFO falls completely outside the scope of this paper. However, in the sequel we give a brief explanation of the elements of this ontology which will be used in the instances of the modeling profile employed in this paper. These elements are summarized in Figure 3.

The model depicted in figure three shows only types of entities, i.e., from a philosophical standpoint, it is an ontology of universals, not one of particulars. A funda-

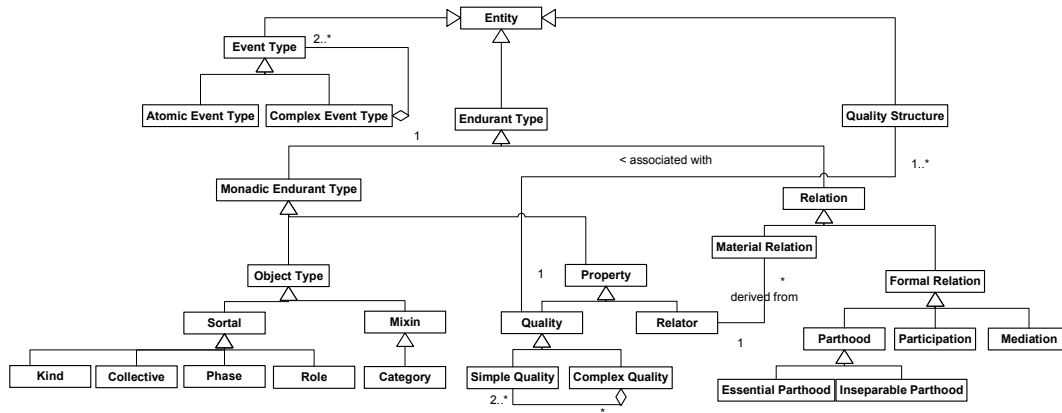


Figure 3. Modeling profile used for developing the ontology

mental distinction in this ontology is the one between *endurants* and *events*, which roughly reflects the common-sense distinction between objects (e.g., a car, a person) and an event (e.g., a race, a business transaction). Endurants persist in time maintaining their identity. Events, in contrast, unfolding in time with their multiple temporal parts and can be either *atomic* or *complex*, the latter which is composed of (possibly complex) events. A formal relation of *participation* is defined between endurants and events. Endurant types can be either types of *monadic* entities or *relations*. Monadic entities, in turn, can be further categorized into *Objects* (again, the stereotypical examples in natural language) and *Properties*. Property instances are entities which are existentially dependent on other entities, in the way, for example, that the color of an apple or the charge of a conductor depends on the apple (conductor) in order to exist. Property instances can be single entities (*qualities*) or multiple entities (*relators*). Examples of the latter include a marriage, a covalent bond, an employment, an enrollment. These entities are existentially dependent on multiple entities and, for this reason, provide the material connection between these entities. In other words, we can say that they are the foundation for *material relations* such as *being married to*, *being connected to*, *working at*, *studying at*, etc. Thus, material relations require relators in order to be established. *Formal relations*, in contrast, hold directly between individuals. In this paper, we consider the formal relations of *parthood*, *participation* and *mediation* (a type of existential dependence). For the relation of parthood we further recognize the case of *essential parthood* – in which a whole cannot exist without that specific part; and the case of *inseparable parthood* – in which a part cannot exist without that specific whole. Qualities can be either *simple* (e.g., weight, age, temperature) or *complex* (e.g., color). Every quality type is associated to a *quality structure*, which can be understood as a measurement structure (or a space of values) in which individual qualities can take their values. For instance, the quality type weight is associated to a space of values which is a liner structure isomorphic to the positive half-line of the real numbers. The color quality type instead is associated to a tridimensional quality structure composed of the dimensions of hue, saturation and brightness. Therefore, every individual color takes its value as a point in this tridimensional conceptual space. Finally, while persisting in time, objects can instantiate several object types. Some of these types an object instantiates necessarily (i.e., in every possible situation) and it defines (from a metaphysical standpoint) what the object is. These are the types named *kind* (for general

objects) and *Collective* (for collection of entities). There are however types that an object instantiates in some circumstances but not in other circumstances. These are named *phases* and *roles*. Whilst phase is a type instantiated in given time period but not necessarily in all periods, a role is a type instantiated in a given context such as the context of a given event participation or a given relation. Finally, a category is a type that classifies entities that belong to different kinds but that share a common essential property (i.e., a property that they must not lack).

Based on this approach, the ECG ontology is composed by (i) structural conceptual models, (ii) FOL axioms and (iii) a terms' dictionary. The next section presents (i) and (ii); for brevity, however, we do not present (iii) in this paper.

4. The Electrocardiogram Ontology

Considering that the main goal of this ontology is to bring in an electrocardiogram theory which is independent of specific applications, the competency questions we have defined reflect this intent. In this way, they lead to a mapping between heart activity and electrocardiography concepts as follows.

- CQ1. What conditions must be satisfied for the heart to play the role of a blood pump?
- CQ2. What conditions must be satisfied for the heart is able to pump blood to both systemic and pulmonary circulation?
- CQ3. What is in the background of an ECG recording session?
- CQ4. What is the source of an ECG record?
- CQ5. How can one obtain the ECG records acquired in the scope of one treatment?
- CQ6. How does an ECG recording device acquire an ECG record?
- CQ7. What does the P wave represent in the ECG waveform?
- CQ8. What does the QRS complex represent in the ECG waveform?
- CQ9. What kind of information does a physician use to identify variations in the morphology and timing of events in the ECG for inferring an interpretation?

For addressing those questions, we have developed the following sub-ontologies: (i) heart, (ii) bioelectric phenomena, (iii) circulatory phenomenon, (iv) ECG human protocol, and, lastly, (v) Electrocardiography (ECG²). These ontologies complement each other to constitute the electrocardiogram (ECG) ontology (Figure 4). They are connected by relations between their concepts as well as by formal axioms. These axioms answer the competency questions introduced above in order to allow: (i) a rich

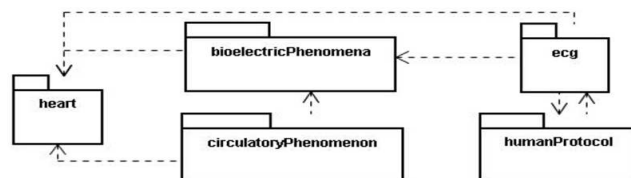


Figure 4. ECG ontology overview

² ECG also stands for Electrocardiography. In this paper, however, the usage context makes unambiguous whether ECG is used for expressing the electrocardiogram or the Electrocardiography.

expressivity that can not be reached only by the graphical model; (ii) inferences (by the ontology codification); (iii) an evaluation of the truthfulness with the ontology purpose; and (iv) identifying inconsistency. The next subsections elaborate on the sub-ontologies. The stereotypes of the modeling profile used are rendered in italics.

4.1. Heart Sub-ontology

This sub-ontology captures the structure of the human heart at a high abstraction level, see Figure 5. As shown in this figure, for a given entity x , the UML ontological profile employed here makes explicit the distinction between a type that defines what an entity is in the ontological sense (e.g., the *Kind* heart) from types which merely describes properties that x shares with other entities of different kinds (e.g., the *categories* chamber and muscle). These distinctions among the categories of object types, renders the language a much higher expressive power than languages such as standard UML, EER or OWL typically used for conceptual modeling and domain ontology engineering.

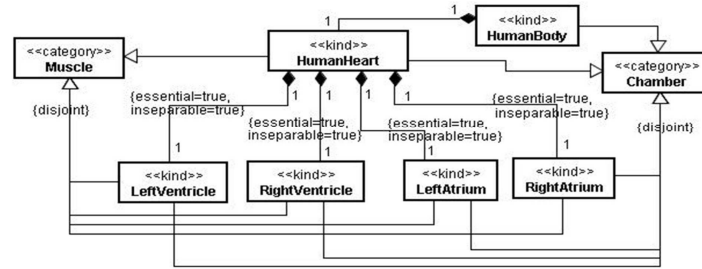


Figure 5. The heart sub-ontology

A human heart has as essential parts left and right atriums and ventricles (represented by the *essential=true* tagged value in the part-whole relation notion). Both atriums and ventricles are also muscles and chambers and as *kinds* of entities, they are disjoint to each other. Finally, atriums and ventricles do not exist apart from the heart as a whole, i.e., they are inseparable parts of the heart they compose (represented by the *inseparable=true* tagged value).

4.2. Bioelectric Phenomena Sub-ontology

Following the textual description laid up in Section 2, one may notice that the heart pumping activity is a natural consequence of bioelectric phenomena generated at special heart cells named pacemaker cells. The SA node and the AV node both constitute the main regions of the myocardial muscle where these cells spontaneously give rise to an electrical impulse generated by its excitation, see Figure 6.

The atrium contractions in answer to the SA electrical impulse characterize material relations between the SA impulse and both atriums' *role* as a pump. These contractions are labeled with the *relator* stereotype. During these contractions the SA node electrical impulse is also conducted to the AV node. At this place, it overdrives the AV node impulse. The His-Purkinje electrical impulse then takes place as a composition of the SA impulse and the AV impulse. It has a *phase* on the bundle of His and a *phase* on the Purkinje fibers. The latter is responsible to left and right ventricles' contractions, characterizing this way, the ventricles' *role* as a pump (Figure 6).

The axioms A1 and A2 formalize the characterization of the ventricles as pumps. The

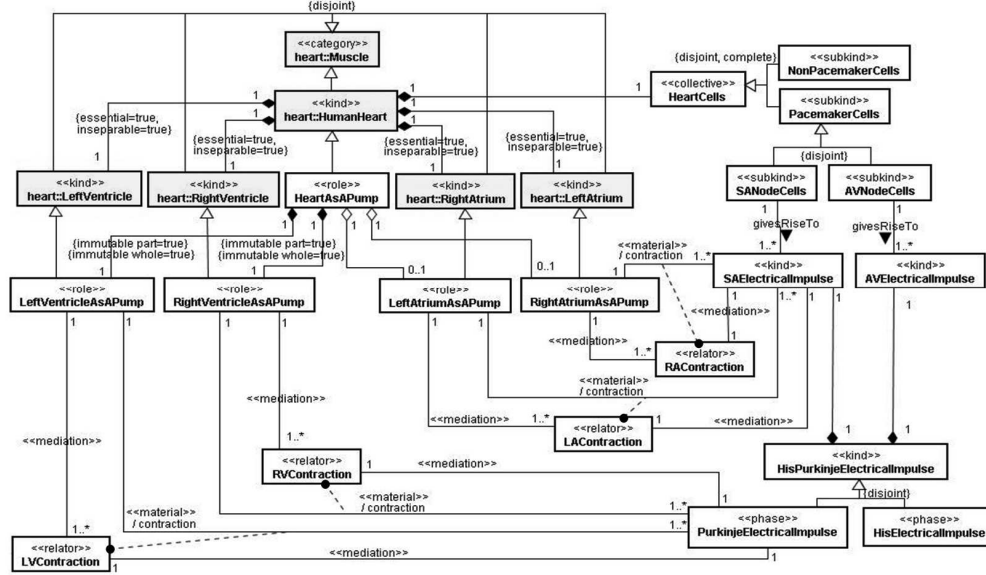


Figure 6. Bioelectric phenomena sub-ontology

- (A1) $\forall x (\text{leftVentricleAsAPump}(x) \rightarrow \exists y,z (\text{purkinjeElectricalImpulse}(y) \wedge \text{lvContraction}(z) \wedge \text{mediates}(z,x) \wedge \text{mediates}(z,y)))$
- (A2) $\forall x (\text{rightVentricleAsAPump}(x) \rightarrow \exists y,z (\text{purkinjeElectricalImpulse}(y) \wedge \text{rvContraction}(z) \wedge \text{mediates}(z,x) \wedge \text{mediates}(z,y)))$
- (A3) $\forall x,y (\text{leftVentricleAsAPump}(x) \wedge \text{rightVentricleAsAPump}(y) \rightarrow \exists z \text{ heartAsPump}(z) \wedge \text{partOf}(x,z) \wedge \text{partOf}(y,z))$

atriums' roles as pumps do not interfere in the heart's role as a pump. The ventricles as pumps, rather, constitute the heart's role as a pump. This is captured by A3 as follows. These three axioms answer the CQ1 competency question.

4.3. Circulatory Phenomenon Sub-ontology

As a consequence of the contractions performed by both atriums and ventricles, the heart as a pump is responsible for the blood circulation around the human body. The lungs and peripherals (i.e., head, torso and limbs) are chambers as well as the atriums and ventricles. The heart moves blood from its chambers to the peripherals and to the lungs as well as from the peripherals and the lungs to its chambers (see Figure 7).

With the heart working as a pump, the whole blood in the human body moves from one of such chambers to another. It is difficult, however, to define an identity criterion to discern one "blood portion" entity in the human body. That is because a blood portion contained in one chamber at one time point will probably change all its properties (e.g. volume, oxygen level, etc.) and container (chamber) on the next time point. Thus, the problem of homeomerosity of quantities addressed in [Guizzardi 2005] is further complicated in the case of blood circulation. It is important to emphasize that the models produced here, i.e., domain ontologies, are intended to focus on structural aspects of the domain (as opposed to dynamic ones). For this reason, the representation of the blood circulation, which is an intrinsically dynamic phenomenon requires a non-trivial mode of representation, namely, the representation of blood portions (the entity

located in each specific chamber at each specific timepoint) as snapshot entities that compose the entire blood contained in a human body. In this way, a human body plays the *role* of a container since it contains a part of the blood in human body. Therefore, the blood storing *relator* concept characterizes a material relation between the *role* of a body chamber as a blood container and a part of blood in the human body (Figure 7).

Figure 7. Circulatory phenomenon sub-ontology

The axioms A4, A5 and A6 state the conditions related to the heart as a blood pump. Once the heart is able to work as a pump (see Subsection 4.2), it is then able to realize its blood pumping function. Nevertheless, for this function to be actualized there must be blood in both of its ventricles, as A4 and A5 axioms state. The heart then may pump blood as it is conveyed by A6. These axioms address CQ2.

The human protocol drives the development of ECG recording sessions as follows (Figure 8). An ECG recording session is a *complex event* that may take place if the following conditions are met: (i) there is an ECG recording device which *participates* in it, (ii) there is a person that *participates* in it playing the role of patient, and (iii) there is a physician who is also a person that has requested it in the scope of a treatment of the patient (see axiom A7). Thus, the treatment *relator* mediates a patient and a physician by means of a material relation.

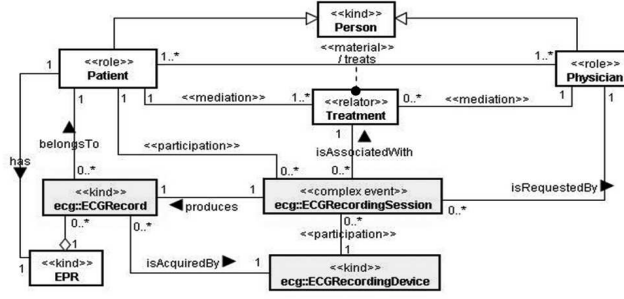


Figure 8. Human protocol sub-ontology

- (A7) $\forall x (ecgRecordingSession(x) \rightarrow \exists !rd, pt, ph, tr (ecgRecordingDevice(rd) \wedge patient(pt) \wedge physician(ph) \wedge treatment(tr) \wedge (participates(rd, x) \wedge participates(pt, x) \wedge isRequestedBy(x, ph)) \wedge (isAssociatedWith(x, tr) \wedge mediates(tr, ph) \wedge mediates(tr, pt)))$
- (A8) $\forall x (ecgRecord(x) \rightarrow \exists !rs, rd, pt (ecgRecordingSession(rs) \wedge ecgRecordingDevice(rd) \wedge patient(pt)) \wedge (produces(rs, x) \wedge isAcquiredBy(x, rd) \wedge belongsTo(x, pt))$
- (A9) $\forall x, y (treatment(x) \wedge ecgRecord(y)) \rightarrow (producedInScopeOf(y, x) \leftrightarrow \exists z ecgRecordingSession(z) \wedge isAssociatedWith(z, x) \wedge produces(z, y))$

patient who was the subject of the session (see axiom A8). The ECG records of one patient may be aggregated in the electronic patient record (EPR) of the patient. Furthermore, a physician can see all the ECG records produced in the scope of one treatment as it is captured by the axiom A9. In other words, formula A9 states that an ECG record y is produced in the scope of treatment x iff y is produced by a recording session z associated with x . The competency questions CQ3, CQ4 and CQ5 are addressed by the axioms A7, A8 and A9 respectively.

4.5. Electrocardiography (ECG) Sub-ontology

This sub-ontology conveys the concepts and relations of Electrocardiography itself. In doing so, it uses some concepts presented in the sub-ontologies previously discussed. As Figure 9 shows, an ECG recording session has a start time and date as well as time duration. Following the methodological guidelines in [Guizzardi 2005], these properties are represented by explicitly representing the conceptual spaces in which they are measured, i.e., in which they can take their values. These conceptual spaces can be either one-dimensional (e.g., sample rate, cardiac frequency) or multi-dimensional (e.g., time, date) and are represented in this modeling profile as simple and structured *datatypes*, respectively.

An ECG recording device can measure the heart activity of a patient by means of electrodes. These electrodes are placed at the body surface of the patient accordingly to an ECG lead (see Section 2). This lead characterizes a material relation between the patient's body and one or two electrodes. Once this relation is established, the electrode(s) and the patient's body are ready for participating in an observation (*an atomic event*). This event occurs at one time point and gives rise to a sample, see axiom A10. The sample concept represents the record of the electrical potential value (in mV) taken by the ECG recording device through one observation. In an ECG recording session, however, the recording device performs multiple observations evenly spaced in time, i.e., an observation series (*a complex event*). These multiple observations then give

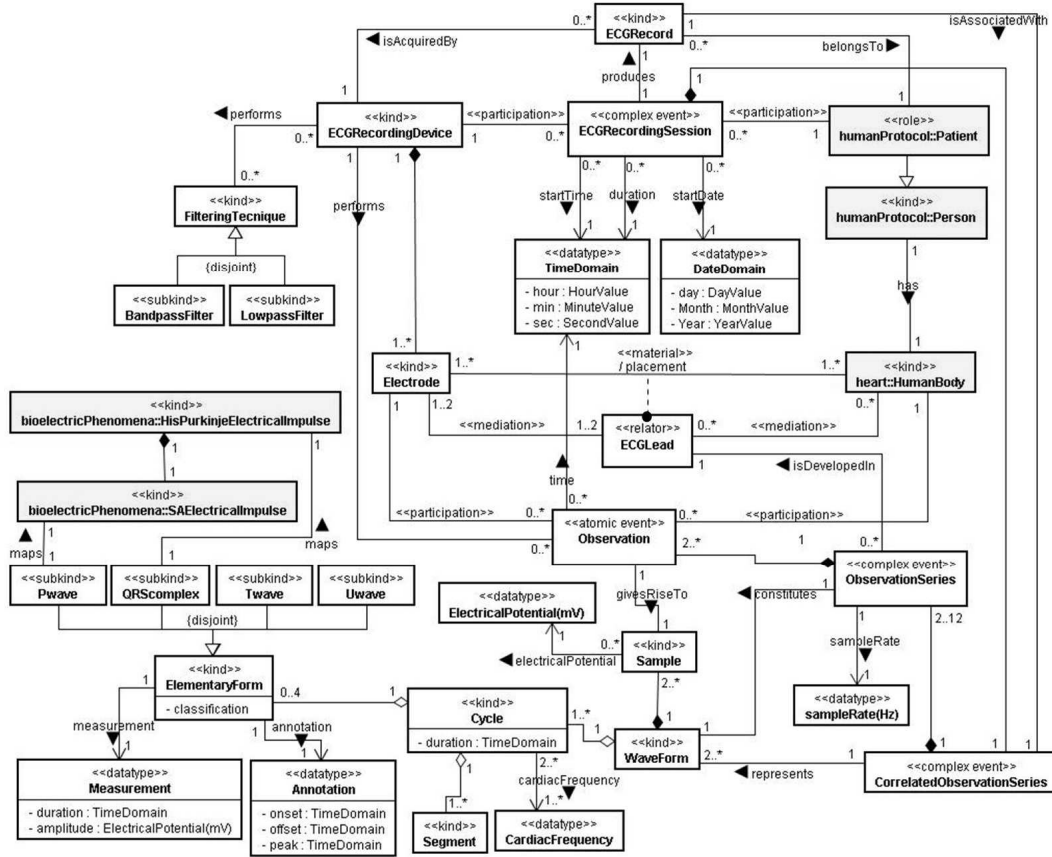


Figure 9. Electrocardiography (ECG) ontology

(A10) $\forall x (\text{observation}(x) \rightarrow \exists et, \exists !rd, ld, hb, sp, tm (\text{ecgRecordingDevice}(rd) \wedge \text{performs}(rd, x) \wedge \text{electrode}(et) \wedge \text{isPartOf}(et, rd) \wedge \text{ecgLead}(ld) \wedge \text{mediates}(ld, et) \wedge \text{humanBody}(hb) \wedge \text{mediates}(ld, hb) \wedge \text{participates}(et, x) \wedge \text{participates}(hb, x) \wedge \text{sample}(sp) \wedge \text{givesRiseTo}(x, sp) \wedge \text{time}(tm) \wedge \text{isProprietyOf}(tm, x)))$

(A11) $\forall x (\text{pWave}(x) \rightarrow \exists !y (\text{saElectricalImpulse}(y) \wedge \text{maps}(x, y)))$

(A12) $\forall x (\text{qrsComplex}(x) \rightarrow \exists !y (\text{hisPurkinjeElectricalImpulse}(y) \wedge \text{maps}(x, y)))$

(A13) $\forall x (\text{elementaryForm}(x) \rightarrow \exists !y, z (\text{measurement}(y) \wedge \text{annotation}(z) \wedge \text{isPropertyOf}(y, x) \wedge \text{isPropertyOf}(z, x)))$

rise to multiple samples which are taken at one sample rate. As a result, an observation series constitutes a waveform which is composed by such samples. In addition, considering the different observation series which are taken at the same time at different ECG leads, we have a correlated observation series (another *complex event*). The result of this correlated observation series, i.e., a set of correlated waveforms, is the basic information laid up in an ECG record.

A waveform of the heart activity, as a rule, has multiple cycles. Each cycle represents each heart beat. We can identify in a cycle both elementary forms and segments that compose the cycle, see Figure 2-(i). The elementary forms are the main characteristics of an electrocardiogram. Both morphology and duration of such waves are full of meaning to the reader of an ECG. That is what we have mapped through the sub-ontologies presented in this paper. The main elementary forms are P wave and QRS

complex. The former maps the SA electrical impulse that is responsible for the atriums' contraction, see axiom A11. The latter maps the His-Purkinje electrical impulse that is responsible for the ventricles' contraction, see axiom A12. By putting together all the perspectives of the heart activity presented in this paper, we can notice the meaning of P wave and QRS complex elementary forms. An elementary form may be annotated as well as it may be measured. These two indications constitute complex properties of an elementary form. They are related to the morphology and the timing of events (in the waveform) that provide meaning to the physicians, see axiom A13.

The segments of a cycle only have meaning w.r.t. the duration of events. This duration can be derived from the annotations of the elementary forms. Over multiple cycles, it is possible to obtain the cardiac frequency, i.e., a number of heart beats per time. This information is paramount to the physicians' analysis. The axioms A10, A11, A12 and A13 answer the questions CQ6, CQ7, CQ8 and CQ9 respectively. Thus all the nine CQ's introduced are answered by the axioms presented in this paper.

5. Related Work

There are quite a few research initiatives committed with ontology-based model and formalization of the ECG. Nonetheless, as a rule such initiatives are committed to the use of technologies and tools such as OWL and Protégé (e.g. [Kokkinaki et al. 2005]) and, consequently, are obliged to rely on models of low expressivity. Moreover, they are remarked with specific purposes w.r.t. specific applications in information systems that restrict their conceptualizations. Finally, there are initiatives aimed to promote interoperability among ECG standards (see Section 1). An example is [Orgun 2003], an effort for allowing the exchange of HL7 messages between heterogeneous healthcare databases. However, as previously discussed, these initiatives tend to be too much driven by both requirements of specific applications and less committed to maximizing expressivity, clarity and truthfulness w.r.t. the domain.

6. Conclusions

This work has presented an ontology of the electrocardiogram domain. This ontology focuses on structural concepts existing in this domain, leaving dynamic concerns as a matter of future work. We advocate that the theory of the ECG conveyed in this paper provides (i) a knowledge repository about this universe of discourse; (ii) a reference conceptual model that may be used for promoting interoperability in Health Informatics; and (iii) a usefulness formal model for ECG analysis systems. These contributions cover a gap in literature w.r.t. ontological approaches for modeling biomedical data. However, there is a need for future research on suitable languages to model biophysics phenomena such as the dynamics of the heart activity. Such enhanced languages may improve the results of ECG analysis systems.

Additional future work includes: (i) the design and implementation of the ECG ontology in one or more codification languages (e.g., OWL, F-Logic) as a proof of concept; (ii) a concrete case study of the use of ontology for interoperating existing ECG standards (e.g., HL7, FDADF, SCP-ECG) and (iii) the extension of the ontology for covering as far as possible pathological situations.

7. Acknowledgments

The first author would like to thank Marcelle Olivier for fruitful discussions and for providing valuable input to the issues of this paper. This research has received financial support from FAPES (grant no. 31024866/2005) and CNPq (grant no. 50.6284/04-2).

8. References

- Ackerman, M., et al. (2002). "State-of-the-Art Telemedicine/Telehealth: An Internal Perspective". Proceedings of a symposium in *Telemedicine Journal and e-Health*, Chapter 6: Telemedicine Technology. Ann Arbor, Michigan, USA.
- Andreão, R., Pereira Filho, J., Calvi, C. (2006). "TeleCardio - Telecardiology for Patient Support in Hospital and Domicile Environments" (in Portuguese). In *X Brazilian Conference in Health Informatics*, Florianópolis, Brazil.
- Brown, B., et al. (2002). "FDA XML Data Format Design Specification". Retrieved July 9, 2007, from <http://xml.coverpages.org/FDA-EGC-XMLDataFormat-C.pdf>.
- Falbo, R. A. (2004). "Experiences in Using a Method for Building Domain Ontologies". In Proc. of 16th Conf. On Software Engineering and Knowledge Engineering (SEKE'04), pp. 474-477, Banff, Alberta, Canada.
- ELC (2007). ECG Learning Center. <http://www.ecglibrary.com/>.
- Geselowitz, D. (1989). "On the Theory of the Electrocardiogram". In *Proceedings of the IEEE*, v. 77, no 6, pp. 857-876.
- Goldberger A., et. al. (2000). "PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals". In *Circulation* 101(23).
- Guizzardi, G. and Wagner, G. (2005). "Towards Ontological Foundations for Agent Modeling Concepts using UFO". In Agent-Oriented Information Systems (AOIS).
- Guizzardi, G. (2005). "Ontological Foundations for Structural Conceptual Models". Ph.D. Thesis, University of Twente, The Netherlands.
- Guizzardi, G. (2006). "The Role of Foundational Ontology for Conceptual Modeling and Domain Ontology Representation". In 7th International Baltic Conf. on Databases and Information Systems, Vilnius, Lithuania.
- Guizzardi, G. (2007). "On Ontology, ontologies, Conceptualizations, Modeling Languages, and (Meta)Models". In *Frontiers in Artificial Intelligence and Applications, Databases and Information Systems IV*, ISBN 978-1-58603-640-8, IOS Press, Amsterdam.
- Guyton, A. and Hall, J. (1996). *Textbook of Medical Physiology*. Ed. WB Saunders, 1996:803–813. Philadelphia, 9th edition.
- HL7 (2003). "HL7 ECG Annotation Message v3". Retrieved July 9, 2007, from <http://www.hl7.org/V3AnnECG/index.htm>.
- Kokkinaki, A., et al. (2005). "A Semantically Enhanced Environment for SCP-ECG Records". In 3rd European Medical and Biological Eng. Conf. (EMBEC'05), Prague, Czech Republic.
- Orgun, B. (2003). "Interoperability in Heterogeneous Medical Information Systems Using Smart Mobile Agents and HL7". M.Sc. Thesis, Macquarie University, Sydney, Australia.
- Salvador, C., et al. (2005). "Airmed-Cardio: A GSM and Internet Services-Based System for Out-of-Hospital Follow-up of Cardiac Patients". In *IEEE Transactions on Information Technology in Biomedicine*, vol. 9, n1.
- SCP (2002). "SCP-Computer-assisted electrocardiography N02-15". CEN/TC251 Secretariat. Retrieved August 8, 2006, from <http://www.centc251.org/>.
- Wang, H., et al. (2003). "A markup language for electrocardiogram data acquisition and analysis (ecgML)". In *BMC Medical Informatics and Decision Making* 3:4.