

BalOnSe: Temporal Aspects of Dance Movement and Its Ontological Representation

Katerina El Raheb^{1,2(✉)}, Theofilos Mailis^{1,2}, Vladislav Ryzhikov³,
Nicolas Papapetrou¹, and Yannis Ioannidis^{1,2}

¹ University of Athens, Athens, Greece
kelraheb@di.uoa.gr

² Athena Research Center, Athens, Greece

³ Free University of Bozen-Bolzano, Bolzano, Italy

Abstract. In this paper, we propose an approach to describe the temporal aspects of ontological representation of dance movement. By nature, human movement consists of complex combinations of spatiotemporal events, a fact that creates a big challenge for representing, searching, and reasoning about movement-related content, such as movement annotations on video dances. We have defined MoveOnto, a movement ontology whose expressive power captures movements that range from body states and transitions based on the semantics of Labanotation, to generic actions or specialized vocabularies of specific dance genres, e.g., ballet or folk. We combine the ontology description with temporal reasoning in Datalog-MTL, based on temporal rules of the movement events. Finally, we present the specifications and requirements for dance exploration from a user’s perspective and describe the architecture of BalOnSe, a specific system that is currently under implementation on top of MoveOnto according to them. BalOnSe consists of a web-based application with semantic annotation, search, and browsing on the movements, as well as a backend with archival and query processing functionality based on temporal rules.

1 Introduction

Dance multimedia data such as videos can be found in large volumes in multimedia channels, but also in dedicated archives and collections. Various recent efforts are aiming at developing and bringing state-of the art information technologies to the area of digitisation, archiving and preservation of intangible cultural heritage and performing-art content [1–3, 7], as well as investigating bodily knowledge and widening the access to such content through enhanced experiences in various contexts, e.g., education. The question is how to facilitate *findability* and provide usable ways to access, search and browse multimedia content, based on the information related to the movement itself through concepts that are interesting for the users. This information could vary from generic concepts regarding actions e.g., step, turn, jump to specific terminology of dance syllabi e.g., pirouette, arabesque, as well as any verbal descriptions that convey something about the form or the quality of movements.

Semantic information about movements in a video can be added in two ways: automated extractions of patterns and movement recognition and manual annotation of experts. In addition to the above, explicit information can be inferred by reasoning if we add to the system rules to represent common and expert knowledge about movement.

The semantic representation of movement includes many challenges. We describe these challenges and how we address them in our work.

Complexity of the human body and its movement: The human body can create endless different combinations depending on the context. These can vary from functional everyday movements and the execution of specific actions, to various dance techniques, expressive gestures, and non-verbal communication. To address this challenge, we have adopted a modular approach based on theoretical basis on movement analysis such as Labanotation and chorological methodologies. In particular, we have developed a core ontology based on MoveOnto, and we develop specific ontologies to describe the terminology of syllabi of different dance genres. We focus on dance genres that allow this modular approach, since they have specific well-defined vocabularies, and structured rules. The entities of the different ontologies (core and specific) are linked through rules which reflect basic and common knowledge about the movement according to technique and the dedicated literature [16, 20].

Temporal aspects: Movement descriptions imply a temporal description. Motion, as well as stillness, in dance is always connected with a duration, and depending on the orchestration and coordination of the different body parts, the temporal aspects might vary in complexity. This complexity is again related to the dance style and technique and the level of detail which is needed depending on the context. We address this challenge by using DatalogMTL [9], a language for representing temporal ontologies. The language allows both temporal representation and reasoning.

Segmentation and Discretization: Semantic descriptions, the construction of concepts and entities, require clear time and space bounds to identify movement entities. Nevertheless, the human motoric usually includes complex synchronizations and coordination of different body parts that occur in a continuous manner. For example, at the semantic level walking is considered to be a continuous sequence of altering steps from one step to another. On the other hand, if we need to analyze and observe movement there is not always an absolute measure to tell us when exactly the first step ends and when the second step ends. We address this challenge by providing a usable interface where expert users can add annotations by choosing terms from the ontology according to their knowledge and perspective. The terms in the ontology vary from more complex recognized sequences to very simple movements. These annotations are stored in the system for further analysis. We base the segmentation, on the notions of *kinemes*, *morphokinemes* and *motifs* as structural, recognisable components of a dance genre, as have been introduced by Kaepller [17] in an analogous way that phonemes, morphemes and words are the morphological components of language.

To put our ideas into practice we have developed *BalOnSe*, a system that allows to add semantically rich temporal annotations about movement on multimedia content, *store* these annotations, and *search* the content based on the annotated metadata. The *BalOnSe* platform is build upon the *MoveOnto* ontology, a movement ontology whose expressive power captures movements that range from body states and transitions based on the semantics of Labanotation, to generic actions or specialized vocabularies of specific dance genres, e.g., ballet or folk. We combine the ontology description with temporal reasoning in DatalogMTL, based on temporal rules of the movement events. In this paper we describe the theoretical basis and the architecture of our implementation, focusing on the representation and management of the temporal aspects, a feature that was not included in the previous version of the system [15].

2 Requirements and Specifications

The architecture (Sect. 4) of the BalOnSe platform is designed to allow modularity and extensibility, i.e., the integration of different rules and ontologies, including more details about the movement, terminology of other dance genres, context of the performance or the recording. To sum up, the current version of BalOnSe platform provides the following characteristics: (i) Annotation with archival functionalities and user moderation; (ii) Semantic-domain specific search (search the content based on movement and dance terminology); (iii) Rich predefined vocabularies for key-word search; (iv) Usable user interface including video streaming, and preview of statistics of annotations; (v) Temporal Reasoning; (vi) Modularity and extensibility.

Through the BalOnSe interface, the user may (i) employ ontological assertions to annotate a dance video recording using a rich ontology that contains both temporal and non-temporal predicates; (ii) browse through dance-multimedia content by querying the ontological video annotations; (iii) preview the rich multimedia content that resides in the BalOnSe deposit. A key characteristic of our system is that it allows further development and integration of other dance-specific ontologies in order to describe different dance genres. By extending our core ontology, that is designed in order to describe sequences of movements, we have built dance-specific vocabularies for different types of genres, i.e. ballet. Thus, a user may query a ballet repository using movement sequences that are defined in the ballet syllabus.

To what follows, we outline the different technical functional and non-functional specifications resulting from the aforementioned scenarios, based also on similar approaches [18].

Integrated Data Access: One of the features of the BalOnSe platform, is its ability to incorporate annotations that have been already made on dance-performance content. These annotations may be stored in a relational database endpoint and can be easily accessed. This is due to the datalogMTL translator that follows an Ontology Based Data Access (OBDA) [21] approach as a means to enhance end-user direct data access to relational databases.

Utilising Implicit Information: In databases it is typically assumed that only explicit data matters, i.e. the explicit data that are stored in the system. Nevertheless, our system aims at inferring additional information by utilising the initial annotations and some domain-specific knowledge related to dance and movement analysis. This could be reflected by logical formalism, that allows to derive implicit information from the data stated explicitly, typically using some forms of background knowledge, via a reasoning procedure. The use of implicit information can greatly increase the practical benefit of the BalOnSe system.

Temporal Data Processing: Our applications requires for the existence of a temporal component that will allow to correlate annotations representing movement sequences of a dance-recording video. Existing languages, such as the OWL 2 language are not designed to represent temporal information. Thus we should adopt a language that allows for an expressive temporal component. We identify the specific refined characteristics for the underlying temporal language: (i) It should provide for a natural way to annotate movements within a dance video recording of a performance. The annotations of video sequence should refer to an absolute time component, i.e. its event should be mapped to a specific time interval within the video that describes its starting and ending point in seconds or subdivisions of seconds. The user should be able to annotate in absolute time intervals of events within a video sequence. (ii) It should provide for a natural way to describe movement patterns and sequences in terms of recognised movement forms for the dance genre. In contrast to the absolute time that applies when annotating a dance video recording, a movement pattern description, even if expressed in notation, does not depend on absolute time-intervals. Instead, similar to music scores, dance sequence descriptions are expressed on beats, i.e. temporal ratios instead of specific time intervals. This means that a mapping is needed between the time expressed in beats in patters and the absolute time intervals coming from the annotations. (iii) The corresponding language should allow to have a temporal component when describing the background knowledge of our domain. The later will allow to effectively encode complex dance sequences via simple predicates. (iv) Finally the corresponding formalism should allow for an expressive query component. The latter will allow to query our metadata for specific events and the time intervals that these events occurred.

Modular Ontological Representation: Ontology modularisation can be interpreted as decomposing potentially large and monolithic ontologies into (a set of) smaller and interlinked components (modules) [4]. An ontology module can be considered as a loosely coupled and self-contained component of an ontology maintaining relationships to other ontology modules. In our domain of interest, we have the core dance ontology that is used to describe basic human body motion (based on Labanotation) and domain specific ontologies that describe specific dance genres, e.g., ballet. The latter provides users of ontologies with the knowledge they require, reducing the scope as much as possible to what is strictly necessary.

3 Movement Description

The BalOnSe platform allows users to enrich multimedia content such as videos with annotations by choosing specific terms from the MoveOnto ontology. MoveOnto includes terminology of movement on different levels of abstraction. To represent choreographies in an expressive searchable way, we need a strong theoretical basis that allows to describe the required elements of dance (grammar and syntax). The theoretical basis for the MoveOnto representation is based on the Labanotation system for describing movement [16] and Kaepller's choreological approaches that analyse the structural parts of dance in motifs, morphokinemes and kinemes. Under this perspective, the notions of kineme and morphokineme [17], as dance segments are analogous to phoneme and morpheme in morphology linguistics. Kinemes are actions or positions, which have no meaning as units alone, but constitute the building blocks of a dance genre. A morphokine is defined as the smallest kinetic unit that has meaning, where meaning here does not reflect any pictorial or narrative notion, but it is used to indicate movements that are recognized as units from the people practicing a specific dance genre. The third level of dance structure is the motif level. A motif is a frequently occurring combination of morphokinemes that themselves form a short entity.

3.1 MoveOnto

In order to provide a more human readable and computer understandable format, MoveOnto was developed to capture the semantics of Labanotation [12], a notation system for recording and analyzing human movement. It is a symbolic language which allow to create dance notations on paper. The symbols are put in specific columns on the staff which is read from bottom to the top. The vertical axis represents time, while different columns of the staff are dedicated to different body parts. The reader interprets the different symbols based on their shape (movement type, directions, and other symbols), their color (level of movement) and their size (duration). A Labanotation score, as in Fig. 1(a), can be seen as a complex timeline with parallel slots.

The main objective of MoveOnto is the interpretation of symbols into entities that are both human- and machine-understandable. Moreover, in our work, the use of the ontology provides the ability to express complex relationships, restrictions, and rules about the concepts, creating hierarchies and graphs of movement entities and properties, and as a result provides a rich vocabulary for describing dance movements. In fact, in a Labanotation score, since the different movements are represented as events, related to time intervals, all Allen's interval relationships [5] are possible to be found on a dance score, to represent the relative synchronisation of states and transitions of movement.

MoveOnto considers three different levels of describing movement [13, 14]: (*i*) a *Labanotation Concepts* level based on Labanotation; (*ii*) a *Generic Concepts Movement* level that is used to describe terms such as turn, step, slide, and (*iii*) *Specific Vocabulary Concepts* level that is used to represent the terminology related to a specific dance genre, such as the Ballet syllabus.

In this paper we provide some examples of the semantic and temporal relations between the syllabus (specific movements that are meaningful for the genre), and their representation in Labanotation and MoveOnto. In our approach, we create a hierarchy/taxonomy of movements, by classifying them into abstract categories. This hierarchy supports the scalability of the system by giving the opportunity to search movements in different levels of detail. As MoveOnto is based on DatalogMTL, it can express complex inference rules and relationships with a temporal component. Reasoning capabilities support reuse of entities, and allow the system itself to infer explicit knowledge from the stored dance knowledge, e.g., “a Gesture is an Action”.

3.2 Time Representation and DatalogMTL

In our previous approach in order to represent time and synchronisation in the core ontology, we adopted a reification strategy. Time intervals were represented as OWL2 individuals called Temporal Entities [12] and in order to represent the sequential order between time points, the functional property `hasNext` was adopted. Each time interval corresponded to a set of facts representing the dance movements, while the duration was expressed by the datatype property `hasDuration`, relating each time-interval individual to a specific duration.

In this work we propose the use of DatalogMTL to represent our knowledge. DatalogMTL is a language for representing temporal ontologies. The advantages that the DatalogMTL language provides, compared to our previous approach, are the following: *(i)* It provides for a natural way to represent time in dance movements, since it does not demand the introduction of individuals that corresponded to time intervals and a datatype property to represent duration. *(ii)* DatalogMTL allows having a temporal component in complex rules. Thus complicated dance combinations can be effectively encoded via simple predicates. For example, a jeté is a complex ballet jump (Fig. 1(f)) that involves a set of different body movements that take part on time. *(iii)* DatalogMTL provides for an expressive query component, which allows querying the dance ontology for specific events and the time points that these events occurred. This provides for very interesting capabilities since it allows to search for specific movement sequences and find similarities between different types of dance.

3.3 Movement Representations and Rules

In the following section we describe some examples of simple movements, expressed in verbal descriptions, Labanotation, and DatalogMTL and we show the different levels of complexity that might occur in temporal representation. A DatalogMTL *program*, Π , is a set of rules about our domain, and a *data instance*, \mathcal{D} , is a finite set of facts, i.e., in our case video annotations.

The MoveOnto ontology makes the following assumptions: *(i)* Simple and more complex movements are represented as unary predicates that are used to characterise video segments. For example an assertion in \mathcal{D} such as

Left_Leg_Gesture_Middle_Back(: Video1)@[12 s, 13 s]

means that a left-leg-gesture-middle-back movement occurred in Video1 during the [12 s, 13 s] time period. (ii) The time in Π (rules) is represented in beats, in a similar way time is represented in dance notations, and expressed in practice when a dance sequence is taught and analysed. (iii) The time in \mathcal{D} (facts) is represented in seconds of each video file. (iv) In the system, the durational value of the beat is mapped into seconds or subdivisions of seconds for each video file according to its tempo, in a similar way the tempo is given in beats per minute (bpm).

Example 1. In Fig. 1(b) we give an example of an Arabesque Allongé pose, a morphokineme that makes sense for a specific dance genre, in this case ballet. The Arabesque morphokineme, either as a pose or action Fig. 1(d), is a case of describing a single pose or movement with no complex temporal relationship of its integral parts. It consists of different simple kinemes that (i) are happening at the same time period, simultaneously; (ii) share the exact same duration. The DatalogMTL definition of the Arabesque Allongé pose is the following:

$$\begin{aligned} \boxplus_{[0,1]} \text{Arabesque_Allonge}(x) \leftrightarrow & \boxplus_{[0,1]} \text{Right_Support_Midde_Place}(x) \wedge \\ \boxplus_{[0,1]} \text{Left_Leg_Gesture_Middle_Back}(x) \wedge & \boxplus_{[0,1]} \text{Right_Arm_Gesture_Middle_Front}(x) \wedge \\ \boxplus_{[0,1]} \text{Left_Arm_Gesture_Middle_Back}(x) \wedge & \boxplus_{[0,1]} \text{Right_Palm_Facing_Place_Low}(x) \wedge \\ \boxplus_{[0,1]} \text{Left_Palm_Facing_Place_Low}(x) \end{aligned} \quad (1)$$

Based on the previous rule, the occurrence of the separate simultaneous movements (involving the legs, arms, and palms) of an 1 s duration implies an Arabesque Allongé movement of the same duration and vice versa.

Within our knowledge we may have different levels of details.

In Fig. 1(c), a more generic Arabesque is represented where only information about the legs is specified. The more generic movement is represented into MoveOnto by the following DatalogMTL rule:

$$\begin{aligned} \boxplus_{[0,1]} \text{Right_Arabesque}(x) \leftrightarrow & \boxplus_{[0,1]} \text{Right_Support_Midde_Place}(x) \wedge \\ \boxplus_{[0,1]} \text{Left_Leg_Gesture_Back}(x) \end{aligned} \quad (2)$$

while we may also add rules for basic movement details:

$$\text{Left_Leg_Gesture_Back}(x) \leftarrow \text{Left_Leg_Gesture_Middle_Back}(x) \quad (3)$$

DatalogMTL allows to query our knowledge using the different levels of detail. Thus an assertion *Arabesque_Allonge(: Video123)@[23, 24]* in \mathcal{D} , along with the rules in Eqs. 1, 2, and 3 imply that

Right_Support_Midde_Place(: Video123)@[23, 24] and *Right_Arabesque(: Video123)@[23, 24]*.

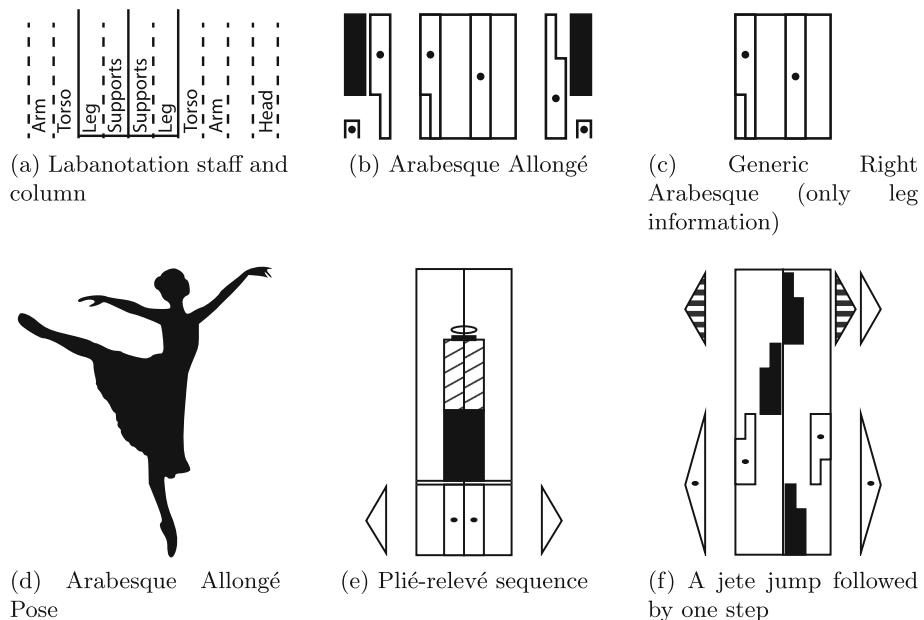


Fig. 1. Labanotation examples

Example 2. In Fig. 1(e) we give an example of a Plié-relevé movement, a more complex morphokineme that includes sequences of simple kinemes: (i) the complex movement is described as a set of sequential movements that are all happening during specific intervals; (ii) all movements overlapping in time, share the exact same duration and are happening simultaneously. In other words, all parallel movements are happening in sequential intervals. These sequential intervals might be equal in time or not.

$$\begin{aligned}
 & \boxplus_{[0,3]} \text{Plie_Releve}(x) \leftrightarrow \boxplus_{[0,1]} \text{Right_Support_Middle_Place}(x) \\
 & \boxplus_{[0,1]} \text{Left_Support_Middle_Place}(x) \boxplus_{[1,2]} \text{Right_Support_Low_Place}(x) \\
 & \quad \boxplus_{[1,2]} \text{Left_Support_Low_Place}(x) \boxplus_{[2,3]} \text{Right_Support_High_Place}(x) \\
 & \quad \quad \quad \boxplus_{[2,3]} \text{Left_Support_High_Place}(x)
 \end{aligned} \tag{4}$$

DatalogMTL allows to infer interesting facts that are happening within a Plié-Relevé-movement interval (similarly with the previous example):

$$\boxplus_{[0,1]} \text{First_Position_Middle}(x) \leftrightarrow \boxplus_{[0,1]} \text{Right_Support_Middle_Place}(x) \wedge \\ \boxplus_{[0,1]} \text{Left_Support_Middle_Place}(x) \quad (5)$$

$$\boxplus_{[0,1]} \text{Plie_in_First_position}(x) \leftrightarrow \boxplus_{[0,1]} \text{Right_Support_Low_Place}(x) \wedge \\ \boxplus_{[0,1]} \text{Left_Support_Low_Place} \quad (6)$$

as well as expressing information that involves sequential movements:

$$\begin{aligned} \boxplus_{[1,2]} \text{Support_Bending}(x) \leftrightarrow \boxplus_{[0,1]} \text{Right_Support_Middle_Place} \wedge \\ \boxplus_{[0,1]} \text{Left_Support_Middle_Place}(x) \wedge \boxplus_{[1,2]} \text{Right_Support_Low_Place}(x) \wedge \\ \boxplus_{[1,2]} \text{Left_Support_Low_Place}(x) \end{aligned} \quad (7)$$

What the above rule says is that if the description starts with an interval where the dancer is in normal position, followed by an interval which defines a low level as a destination, this should allow us to subsume that the dancer has bent his/her knees, so there is an action of bending there. Thus by combining the rules in Eqs. 4 and 7 we infer that the Plié-Relevé movement implies a Support-Bending movement as well.

Example 3. In Fig. 1(f) we give an example of a Jeté jump where parallel and sequential movements occur in combination, thus more complex synchronisation needs to be represented: (i) the complex movement is described as a set of sequential movements that are all happening during specific intervals of the same or different durations; (ii) movements might occur in equal intervals, overlap, start with or end with, so all combinations of intervals relations are possible.

$$\begin{aligned} \boxplus_{[0,4]} \text{Jete_Jump}(x) \boxplus_{[0,1]} \leftrightarrow \text{Right_Step_Low_Place}(x) \wedge \\ \boxplus_{[1,2]} \text{Right_LegGesture_Middle_Back}(x) \wedge \boxplus_{[1,2]} \text{Left_LegGesture_Middle_Forward}(x) \wedge \\ \boxplus_{[2,3]} \text{Left_Step_Middle_Low}(x) \wedge \boxplus_{[3,4]} \text{Right_Step_Middle_Low}(x) \wedge \\ \boxplus_{[0,2]} \text{Right_Arm_Gesture_Middle_Right}(x) \wedge \boxplus_{[0,2]} \text{Left_Arm_Gesture_Middle_Left}(x) \wedge \\ \boxplus_{[3,4]} \text{Right_Arm_Gesture_Middle_Right}(x) \wedge \boxplus_{[3,4]} \text{Left_Arm_Gesture_High_Left}(x) \wedge \\ \boxplus_{[3,4]} \text{Torso_Gesture_High_Right}(x) \end{aligned} \quad (8)$$

4 BalOnSe Platform

BalOnSe is an integrated system that consists of multiple components to support end-to-end annotation and provide access on ontologies about dance. From a user perspective, the BalOnSe system allows to: (i) create temporal ontological annotations of multimedia content; (ii) search by movement predicates; (iii) preview the existing multimedia content that satisfies the queries. Query evaluation is done via the systems query enrichment, unfolding, and execution modules that allow executing complex temporal queries on relational databases. In this section we give some details of the BalOnSe components that address the challenges and specifications described in Sect. 2.

4.1 BalOnSe Components

The application provides for the following components in order to perform the intended actions:

1. *Multimedia Content Annotation:* the application provides for an archival system for the videos, along with the corresponding user interfaces (UIs) that allow users to annotate video content and store the corresponding annotations in a relational database system.

The query formulation, transformation, execution, and video visualisation procedures are performed in a sequence of stages presented in Fig. 2:

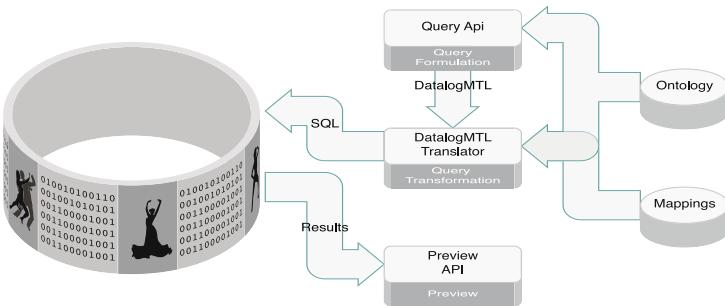


Fig. 2. BalOnSe architecture

2. *Query Formulation*: Data sources can be explored using a query formulation tool that allows the user to navigate through the ontology and the corresponding temporal component. The current version of our query formulation tool is based on DatalogMTL queries. Future work involves the adoption of faceted browsing techniques that allow end-users to explore a collection of information by applying multiple filters.
3. *Query Transformation*: The aforementioned expressions are sent to the corresponding query transformation engine for processing. The processing includes rewriting against the ontology and further unfolding into relational SQL queries based on the corresponding mappings [9].
4. *Query Execution*: SQL queries are executed by a relational database management system.
5. *Preview*: The results of query execution are video fragments corresponding to the complex movements that were queried. The user is able to preview the corresponding video fragments.

4.2 User Interface

A key characteristic of the BalOnSe platform is its simplicity for annotating video content and search. The platform's web-based interface is illustrated in Fig. 3 and can be logically divided to 5 basic interfaces. The user is able to perform annotations using labels of predefined vocabularies on the core ontology and morphokinemes on domain-specific ontologies (Fig. 3D). Each annotation has a starting and an ending time point. For formulating queries, we have adopted a transparent approach where the user asks only for specific labels corresponding to the terminology of (morpho)kinemes (Fig. 3B). Finally, the user can directly search for multimedia objects based on their titles (Fig. 3A), and furthermore preview multimedia content (Fig. 3C).

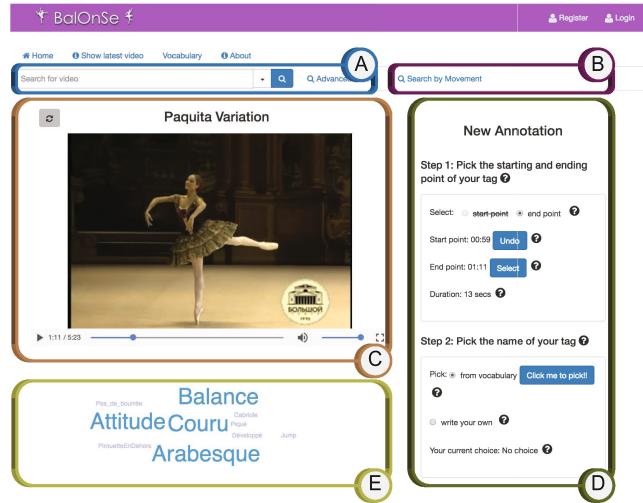


Fig. 3. BalOnSe user interface

4.3 Annotations Storage and Management

The application provides an archival system for the videos while both video metadata and user annotations are stored in a relational database. By taking advantage of the modular characteristics of our ontological representation, the Multimedia Content Annotation Interface allows the user to annotate a dance video either by adopting simple movements (kinemes) on the core ontology, or by using predefined combinations of simple movements (morphokinemes) on the domain-specific ontology (e.g. the ballet ontology).

Example 4. The dance expert may annotate by using the core ontology, that in Video_123 a Right_Support_Middle_Place movement took place during the period [25 s, 26 s] and a Left_LegGesture_Back movement took place during [25 s, 26 s]. Alternatively, the dance expert may assert that a Right_Arabesque movement was performed between [25 s, 26 s]. Based on the definition of Right Arabesque in Eq. 2, the two annotations are equivalent.

Database Schema: We will briefly present the database schema, which implements a basic entity relationship model with three main entities: (**User Accounts**, **Videos**, and **Annotations**). We focus on the **Annotations** table, which stores our temporal information:

Annotations is the table that stores the experts annotations. A new record is saved for each annotation that is chosen by the user for a specific video segment. It has the following attributes: The **VideoId** contains an identifier that is unique for each video file. The **Movement** attribute represents either a kineme in the core ontology, or a morphokineme in some of its extensions. **Comment** is a user defined characterisation of the annotation the user just created. **PostTime** is the time the user made the annotation, this attribute provides the opportunity to

have log history of tags added by the same users in different times. **StartTime** is the time the annotation starts being observed in the video. **EndTime** is the time the annotation stops being observed in the video. **StartTime** and **EndTime** are expressed in seconds of the video time.

Example 5. Based on Example 4 for the core ontology, the user will make the following annotations: a right-support-middle-place movement took place between [25 s, 26 s] of the specific video and a left-leg-gesture-back movement took place between [25 s, 26 s] of the specific video. The two annotations are stored as displayed in Fig. 4.

Annotations							
VideoId	Movement	StartTime	StopTime	PostTime	Comment	UserId	
123	Right_Support_Middle_Place	25 sec	26 sec	12/09/2016 @ 10:54am	-	24	
123	Left_Leg_Gesture_Back	25 sec	26 sec	12/09/2016 @ 10:54am	-	24	

Fig. 4. Annotations in the relational database

Therefore, for each video all the annotations provided by different users are stored, as well as different annotations of the same user. The database stores the multiple tags along with the information of the user that provided them. In the current version, the annotations of users are stored and each time the video appears, a tag cloud shows all existing annotations.

4.4 Query Transformation and Execution

This section is dedicated to the enrichment and unfolding stages that occur during a query execution cycle. The DatalogMTL translator is responsible for transforming DatalogMTL conjunctive queries and their temporal component, to relational SQL queries.

Ontology Mappings: The relationship between the ontological vocabulary and the schema of the data is maintained by a set of mappings. Intuitively, a mapping assertion consists of a source (an SQL query retrieving values from the database) and a target (defining RDF triples with values from the source). The DatalogMTL translator exposes relational databases as virtual datalog assertions by linking the terms (predicates) in the ontology to the data sources through mappings. The W3C standard for expressing customised mappings from relational databases to RDF datasets is R2RML (R2RML: RDB to RDF Mapping Language) [11].

DatalogMTL translator extends R2RML mappings, so that they incorporate both predicates of higher arity, as well as temporal information. For the technical details of these extensions the reader may refer to [9]. We will try to explain with a simple example how R2RML mappings work for our use cases.

Example 6. The ontology instance assertions in Fig. 4 can be populated from a database by means of the following mappings:

```

Right_Support_Midde_Place(:Video_{VideoId})@[StartTime,StopTime]
  ← SELECT VideoId, StartTime,StopTime FROM Annotations
    WHERE Movement='Right_Support_Midde_Place'
Left_Leg_Gesture_Back(:Video_{VideoId})@[StartTime,StopTime]
  ← SELECT VideoId,StartTime, StopTime FROM Annotations
    WHERE Movement='Left_Leg_Gesture_Back'.

```

By combining the mappings and the data in the `Annotations` table (Fig. 4) we will get the facts: `Right_Support_Midde_Place(: Video123)@[25 s, 26 s]` and `Left_Leg_Gesture_Back(: Video123)@[25 s, 26 s]`. This data can then be queried using DatalogMTL, by translating the DatalogMTL queries into SQL queries over the relational databases. This translation process is transparent to the user and performed in the rewriting and unfolding steps.

Rewriting, Unfolding, & Execution: DatalogMTL uses two-step approach to answer the user queries described in Sect. 3: (i) First, an initial DatalogMTL query Q_1 together with the ontology is rewritten into a query Q_2 over the virtual relations that represent temporal data in conceptualised form. Example 6 demonstrated a query that populates one such relation, namely, `Right_Support_Midde_Place`, which has columns `VideoId`, `StartTime`, `StopTime` to represent temporal information. (ii) Q_2 is then combined with all the relevant mappings to obtain the final query Q_3 that can be evaluated over the raw data.

The process to obtain the query Q_3 involves the computation of *temporal joins* and *coalescings*. To give an idea, for two tables T and S with columns `from`, `to` both containing intervals, temporal join is a table with the same columns that contains all the non-empty intervals i , such that $i = j \cap k$, for some j from T and k from S . On the other hand, coalescing, for a table T as above, is a table T' which contains only one “covering” interval $[i_1, j_n]$ for a sequence of (right-)overlapping intervals $[i_1, j_1], \dots, [i_n, j_n]$.

Example 7. A simple DatalogMTL query asking for the videos that contain a Right-Arabesque movement and the time that this movement was performed has the form $Q_1(x, \delta) = Right_Arabesque(x)@\delta$ where x corresponds to the video file and δ corresponds to the time interval that the movement was performed. By taking into account the definition of the right-arabesque movement in Eq. 2 and the mappings in Example 6, the rewriting and unfolding procedure will execute a complicated query Q_3 over the table `Annotations` of the relational database. For space limitations, we won’t provide the full rewriting and unfolding of the initial query, instead, the interested reader may refer to [9] for more information on the corresponding procedures. The query answer will be the tuple $(: Video_123, 25, 26)$ meaning that `Video_123` contains a `Right_Arabesque` during the $[25 \text{ s}, 26 \text{ s}]$ time interval.

5 Novelty and Impact

Related Work: We describe previous advancement in the field of multimedia annotation for performing arts. Bertini et al. [8] have presented an overview of

approaches and algorithms that exploit ontologies to perform semantic video annotation and an adaptation of the First Order Inductive Learner technique to the Semantic Web Rule Language. Ramadoss and Rakummar [22] have presented the system architecture of a manual annotation tool, a semiautomatic authoring tool, and a search engine for the choreographers, dancers, and students of pop Indian dance. Singh et al. [24] presented the Choreographic Notebook, a multimodal annotation tool supporting the use of text and digital ink to be used during the production process of contemporary dance. A similar multimodal annotation approach has been presented by Cabral et al. [10] in the Creation-Tool, while Ribeiro et al. [23] have extended this methodology through addressing the issues of transforming the 2D annotations into 3D visualizations. Another relevant work, in the field of non-verbal communication, is the Anvil interface and the corresponding schema of manual annotation for conversational gestures, which eventually supports the recreation of 2D animation based on time and special descriptions of the gestures on videos [19].

Novelty: In comparison to the existing approaches, BalOnSe is the first platform that allows to represent time within terminological knowledge, allowing to define complex morphokinemes where parallel and sequential movements occur in combination. Moreover, it is the only such platform that uses an OBDA approach, separating the assertional from the terminological knowledge and exploiting database technologies to store and retrieve information.

Impact: The BalOnSe platform is currently in a beta version and its functionality has been presented to dance experts and got positive feedback. In particular, the experts were asked to assess the platform with respect to the specifications introduced in Sect. 2. The experts gave a positive feedback on how the proposed solution addresses the *Integrated Data Access* requirement: the OBDA approach allows to integrate to the ontology video annotations that have already been collected and stored in external data sources such as relational databases. The experts gave positive comments on the platform's ability to derive implicit information using logical reasoning on temporal information, thereby satisfying requirements for *Utilization of Implicit Information* and *Temporal Data Processing*. Finally, the *Modular Ontological Representation* approach that has been adopted by our system, was highly welcomed since it allows to represent information in different levels of abstraction by having the core dance ontology to describe kinemes and domain specific ontologies that introduce morphokinemes for specific dance genres.

6 Conclusions and Future Work

We have proposed an approach to describe the temporal aspects of ontological representation of dance movement. Our approach is based on the MoveOnto ontology that adopts the DatalogMTL language to represent movement in time. To put our ideas into practice we have developed the BalOnSe platform, the first platform that allows to store and query performing-art content containing complex temporal information.

Future work involves extending our platform with faceted browsing techniques, a suitable paradigm for querying ontology repositories [6]. Moreover, we intend to examine DatalogMTL extensions that will allow us to infer conclusions when there is no perfect synchronisation between sequences of movements. Therefore, we want to extend the DatalogMTL semantics with fuzzy values to express the degree of similarity between movements.

Acknowledgements. This paper includes part of the research which is conducted in the framework of the WhoLoDancE H2020 EU funded project.

References

1. The eclap e-library of performing arts. <http://www.eclap.eu/>
2. Itreasures, capturing the intangible. <http://i-treasures.eu>
3. Whole body interaction learning for dance education. www.wholodance.eu
4. Abbes, S.B., Scheuermann, A., Meilender, T., d'Aquin, M.: Characterizing modular ontologies. In: 7th International Conference on Formal Ontologies in Information Systems (2012)
5. Allen, J.F., Hayes, P.J.: A common-sense theory of time. In: IJCAI, vol. 85 (1985)
6. Arenas, M., Grau, B.C., Kharlamov, E., Marcuska, S., Zheleznyakov, D., Jimenez-Ruiz, E.: Semifacet: semantic faceted search over yago. In: 23rd International Conference on World Wide Web (2014)
7. Bellini, P., Nesi, P.: Modeling performing arts metadata and relationships in content service for institutions. *Multimedia Syst.* **21**, 427–449 (2015)
8. Bertini, M., Del Bimbo, A., Serra, G.: Learning ontology rules for semantic video annotation. In: Proceedings of the 2nd ACM Workshop on Multimedia Semantics, pp. 1–8. ACM (2008)
9. Brandt, S., Kalayci, E.G., Kontchakov, R., Ryzhikov, V., Xiao, G., Zakharyashev, M.: Ontology-based data access with a horn fragment of metric temporal logic. In: Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence, February 4–9, 2017, San Francisco, California, USA, pp. 1070–1076 (2017)
10. Cabral, D., Valente, J.G., Aragão, U., Fernandes, C., Correia, N.: Evaluation of a multimodal video annotator for contemporary dance. In: Proceedings of the International Working Conference on Advanced Visual Interfaces, pp. 572–579. ACM (2012)
11. Das, S., Sundara, S., Cyganiak, R.: R2RML: RDB to RDF Mapping Language (2012)
12. El Raheb, K., Ioannidis, Y.: A labanotation based ontology for representing dance movement. In: International Gesture Workshop (2011)
13. El Raheb, K., Ioannidis, Y.: From dance notation to conceptual models: a multi-layer approach. In: International Workshop on Movement and Computing (2014)
14. El Raheb, K., Ioannidis, Y.: Modeling abstractions for dance digital libraries. In: Joint Conference on Digital Libraries (2014)
15. El Raheb, K., Papapetrou, N., Katifori, V., Ioannidis, Y.: Balonse: ballet ontology for annotating and searching video performances. In: 3rd International Symposium on Movement and Computing (2016)
16. Guest, A.H.: Labanotation: The System of Analyzing and Recording Movement. Routledge, New York (2014)

17. Kaepller, A.L.: Dance structures: perspectives on the analysis of human movement. Akad. Kiadó (2007)
18. Kharlamov, E., et al.: How semantic technologies can enhance data access at siemens energy. In: Mika, P., et al. (eds.) ISWC 2014. LNCS, vol. 8796, pp. 601–619. Springer, Cham (2014). doi:[10.1007/978-3-319-11964-9_38](https://doi.org/10.1007/978-3-319-11964-9_38)
19. Kipp, M., Neff, M., Albrecht, I.: An annotation scheme for conversational gestures: How to economically capture timing and form. *Lang. Res. Eval.* **41**(3–4), 325–339 (2007)
20. Miles, A., Grant, G.: The Gail Grant Dictionary of Classical ballet in Labanotation. Dance Notation Bureau, New York (1976)
21. Poggi, A., Lembo, D., Calvanese, D., De Giacomo, G., Lenzerini, M., Rosati, R.: Linking data to ontologies. *J. Data Semant.* **10**, 133–173 (2008)
22. Ramadoss, B., Rajkumar, K.: Semi-automated annotation and retrieval of dance media objects. *Cyber. Syst. Int. J.* **38**(4), 349–379 (2007)
23. Ribeiro, C., Kuffner, R., Fernandes, C., Pereira, J.: 3d annotation in contemporary dance: enhancing the creation-tool video annotator. In: Proceedings of the 3rd International Symposium on Movement and Computing, p. 41. ACM (2016)
24. Singh, V., Latulipe, C., Carroll, E., Lottridge, D.: The choreographer's notebook: a video annotation system for dancers and choreographers. In: Proceedings of the 8th ACM Conference on Creativity and Cognition, pp. 197–206. ACM (2011)