

Introducing a Language for Human Stance Description

Antônio C. Mabiála¹, Antonio W. Sousa², Norton T. Roman¹(✉),
João L. Bernardes Jr.¹, Marcelo M. Antunes³, Enrique M. Ortega³,
Luciano A. Digiampietri¹, Luis M. del Val Cura⁴, Valdinei F. da Silva¹,
and Clodoaldo A.M. Lima¹

¹ University of São Paulo, São Paulo, Brazil
norton@usp.br

² São Paulo Faculty of Technology, São Paulo, Brazil

³ Central Kung Fu Academy, Campinas, Brazil

⁴ Campo Limpo Paulista Faculty, Campo Limpo Paulista, Brazil

Abstract. The ability to automatically determine a human body's sequence of postures during movement has many practical applications, from the evaluation of the performance of physical activity practitioners to the evaluation and design of the user experience in certain systems. Current representations for such postures, however useful, are not capable of capturing all necessary features for a complete description of a human stance, such as the relationship between non-directly connected body parts. In this article, we introduce a mark-up language for body stance and movement description, designed to allow for the unambiguous representation of movement as well as the extraction of relationships between directly and non-directly connected body parts. Along with the language, we also present a computer program, developed to help end users in codifying stance and movement without having to know the language in detail.

1 Introduction

The ability to automatically determine a human body's sequence of postures during movement has many practical applications, from the evaluation of the performance of physical activity practitioners (e.g. [1]) to patients' physical rehabilitation (e.g. [3,4]), posture assessment (e.g. [6]), aiding in learning (e.g. [16]) or in exercising (e.g. [5]) and even for the evaluation and design of the User Experience in certain systems (e.g. [2]). Such tasks, however, depend not only on some existing apparatus for capturing and tracking the position of body parts (such as Microsoft Kinect, for example), but also on some form of representation suitable for future inspection, comparison or specification of detected and recorded postures.

Current representations, such as the Human Mark-up Language¹ and the Human Performance Mark-up Language², for instance, however useful, are not capable of capturing all necessary features for a complete description of a human stance. More specifically, they have no straightforward way to determine the relationship between non-directly connected body parts, such as the position of the left hand related to the right foot, for example, which is paramount to describe some physical activities, such as dancing and practising martial arts.

In this article, we try to fill in this gap, by introducing a mark-up language for the description of human postures, taking into account not only the relationship between directly connected body parts, but also between indirectly connected ones, along with some symmetry relations. This effort is taken as part of a bigger project, called *Education and monitoring of physical activities through artificial intelligence techniques*³, which aims at the teaching and monitoring of physical activities, specially within the practice of martial arts.

Within this scope, this research aims to define requirements for body posture descriptions in an unambiguous way, allowing different descriptions to be evaluated and compared. We then introduce a language designed to fulfil all these requirements. Formatted as an XML file, the language offers both a reference system in conformance to the ISO/IEC FCD 19774:200x Humanoid animation H-Anim standard, and a description based on the identification of different angles between connected parts. Both descriptions co-exist in the XML: while the former is mainly used to unambiguously describe the stances, the later is meant to be used by the end user, given its more intuitive nature.

Along with the language specification, we also developed a computer program to help users describe postures through a graphical interface, where relationships between body parts can be described by selecting the type of relationship from a previously defined list, along with possible angles between connected parts⁴. With this software, we intend both to reduce the cognitive load on the user, by moving the description details away from the user's attention, and to reduce the possibility of errors in the codification of the posture.

Finally, all produced materials, such as the language documentation and end-user software (along with its source code), will be freely available to the community under a creative commons license. The rest of this article is organized as follows. Section 2 introduces some of current movement notation systems, along with a brief discussion about computational models that use them. In the sequence, Sect. 3 describes some related work on mark-up languages for human postures. Section 4, in turn, presents the specification of the language we

¹ <https://www.oasis-open.org/committees/download.php/60/HM.Primary-Base-Spec-1.0.html>.

² <http://www.sisostds.org/StandardsActivities/StudyGroups/HPMLSGHumanPerformanceMarkupLanguage.aspx>.

³ Process N°: 14.1.00923.86.4.

⁴ Notice that the definition of a sequence of connected angles would eventually lead to the definition of an angle between non-directly connected parts.

developed, whose working model is presented in Sect. 5. Finally, Sects. 6 and 7 present our results and a conclusion to this research.

2 Movement Notation Systems

Created by Joan and Rudolph Benesh to record the movements of classical ballet, the Benesh System [14] is committed to the line of execution and the visual result of the movement. Based on a body position that resembles the T-stance, commonly used by game developers and 3D characters animators, the system uses a set of symbols to show the position of the dancer and the posture being performed, along with five parallel lines to determining the height of body parts positions. Within the same framework, Laban’s Movement and Dance Notation [8] uses a finite model of space, where the body moves immersed in a sphere, separated by three perpendicular planes, to record any type of human movement. The system then uses distinct symbols to represent body parts and movement directions, taking into account the sphere in which the model is fit.

Alternatively, the Eshkol-Wachman Notation [14] divide the space with planes, using pair of numbers to represent important points. Also relying on a set of symbols to represent body parts and movement directions, the system was used in computer graphics, architecture and work related to animal behaviour, being also used to describe the Israeli Sign Language and to evaluate the movements of the Tai Chi Chuan Chinese martial art [7]. Within this notation, the body is represented as a “stick-figure”, with segments and points. The segments connect pairs of joints or points, so that every part of the body has a representation in a manuscript defined by five parallel lines. As in the Laban Notation, Eshkol-Wachman also uses a spherical reference system, but with pairs of numbers representing specific positions in space, as opposed to symbols.

Finally, the HamNoSys – Hamburg Notation System [9] – is a notation system designed to support avatar-independent sign language transcriptions, recording only the relevant aspects of language for the correct execution of gestures. The system defines a spacial volume where gestures are performed, taking into account just hand and arm movements, thereby considering the rest of the body as static. This notation has, in turn, served as the basis for SiGML (Signing Gesture Mark-up Language) [13], a mark-up language used to record gestures in sign languages, also being avatar-independent.

Existing computational models usually rely on some of these notations, such as the one presented in [15] which, based in Laban’s Kinesigraphy, deals with gesture acquisition and synthesis. However interesting, such models usually need some previous knowledge and preparation from the person performing the movements, also needing a controlled environment. Additionally, despite the fact that notation systems try to make easier the record and the understanding of gestures by people with proper training, they are still not suitable for non-specialist in the system. Another noteworthy point about them is that they are not interoperable, being necessary some sort of adaptation or conversion from one system to another.

3 Related Work

The proposal of a mark-up language for human postures and expressions is not something new, with many alternatives being proposed over the last few years, each of them with different objectives and scopes. One such example is VHML – Virtual Human Mark-up Language⁵ – which was designed to encompass various aspects of Human-Computer Interaction. Acting like a super-set of languages, VHML paved the way to other more specific languages, such as EML – Emotion Markup Language, FAML – Facial Animation Markup Language and BAML – Body Animation Markup Language, amongst others. However also designed to be expansible and generic, VHML was primarily created to make easier the communication between humans and avatars.

Others languages, such as Human Markup Language [16], for example, propose to encompass cultural aspects and its specification, since high-level languages can show some problems when defining concepts like angry or happy. Just like VHML, HumanML is a high-level language, being used to record the behaviour and expression of virtual agents, especially in their interaction with the user. A different proposal, but with a similar goal, is XSTEP [11], a mark-up language based on the STEP [10] script language. As a language, XSTEP has the interesting feature of allowing other languages to be embedded in it.

Our research differs from these in that we move away from the goal of describing a human being, be it virtual or not, in its entirety, thereby setting aside aspects such as culture and emotion, or the possibility of having them added to the language. Instead, we focus on how to represent body stances and movements, specially for the use in the realm of martial arts. Hence, and as it will be made clearer in the forthcoming sections, we try to fulfil some requirements specific to this task, being able to describe relations between both connected and non-directly connected body parts (which, in our model, can be described as sequences of positions and angles between all connected body parts along the way from one unconnected part to the other).

4 Language Specification

In order for the language to be computationally useful, it should [12]:

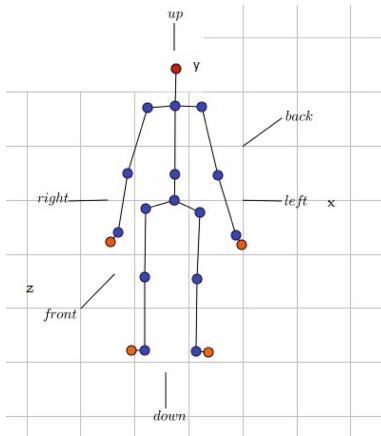
1. Be convenient, in that posture specifications should hide geometrical details away from the user, thereby allowing it to be used in the most natural way possible by domain experts; and
2. Possess a compositional semantics, thereby allowing for complex descriptions to come out as relationships between simpler component descriptions.

Additionally, in our interactions with martial arts experts, we set up a list of further requirements for the language, according to which it should also be able to represent:

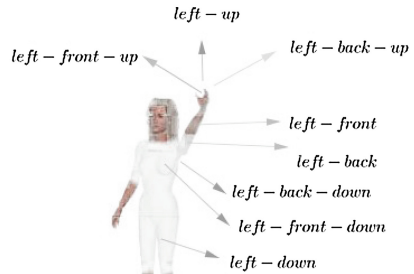
⁵ <http://www.vhml.org/>.

1. The relationship between body parts in symmetry and asymmetry;
2. The relationship between upper body parts and between upper and lower parts, both in the same side and opposite sides of the body;
3. The relationship between the head and the transversal body line (the line crossing the so called Dantian⁶);
4. The relationship between the position of the head in relation to the body balance area, as determined by the position of the feet;
5. The relationship between the head and the body gravity centre; and
6. The relationship between gaze and the direction of subsequent movements.

To comply with these requirements, we begin by defining the starting position of any description, which corresponds to a body standing along the y axis (Fig. 1(a)), with the x axis going from left to right and the z axis moving in the front to back direction. The origin point (0,0,0) is located at the ground level, between the feet. The arms are straight and parallel to the sides of the body, with the palms facing inward towards the thighs. The reference system is builds on these three dimensions. Figure 1(b)⁷ illustrates some combinations based on these three dimensions.



(a) Starting position



(b) Direction references

Fig. 1. Starting position and direction of reference of body parts.

Additionally, the *ISO/IECFCD19774:200x Humanoid animation H-Anim17* specification also contains a set of nodes that are arranged to form a hierarchy. Each node contains other nodes and a number of nodes that specify which vertices correspond to a particular characteristic or configuration.

⁶ Energy centers in Chinese martial arts.

⁷ Adapted from [11].

5 Working Model

The articulated design model consists of points, called elements, which represent different parts of the human body. Based on this representation it is possible to extract angles and establish relationships between them. Figure 2 illustrates the representation parts to be considered in our model, while Fig. 3 shows the working model associated with them. The language, however, was designed to accommodate several elements of human body, not only those described in the working model.

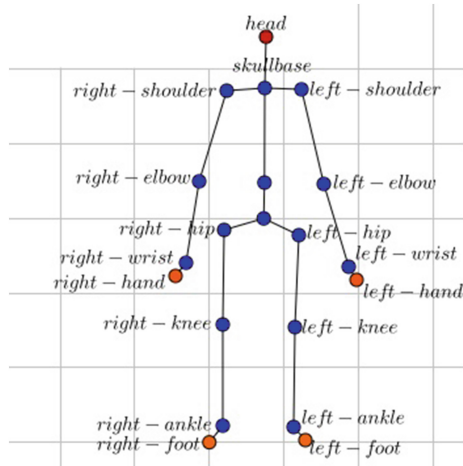


Fig. 2. The human body and its typical joints representation

Since the language was primarily designed for the use in a martial arts context, body parts may assume some pre-defined descriptions, such as the tiger, snake, panther and crane kung-fu stances, for instance. These description apply both to the whole body or to specific body parts, such as the hands, for example. Other elements, not mentioned in the working model, along with other language features, are described in the language documentation⁸.

Following the pivot points described in Fig. 3, base input parameters can be provided to represent stances, thereby ensuring that each of them will have a unique description that can be extracted from the angles of their hinged parts and the relations of symmetry and asymmetry between them. Also, and to ensure more consistency between the elements, we defined a hierarchy between the various body parts, as shown in Fig. 4. Following the hierarchy, in order to describe the angles between the joints of the upper limbs, one has to start from the skull. Similarly, for the angles between the joints of the lower limbs, the hips are the starting point.

⁸ <http://1-dot-term-paper-1121.appspot.com/pack/documentation/index.html>.

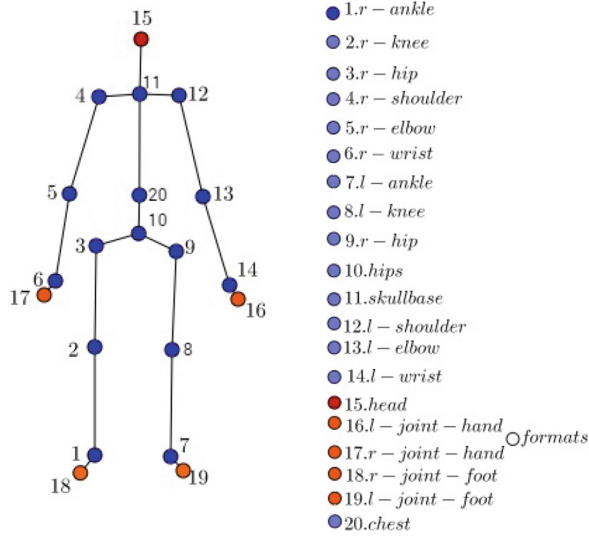


Fig. 3. The working model and its elements.

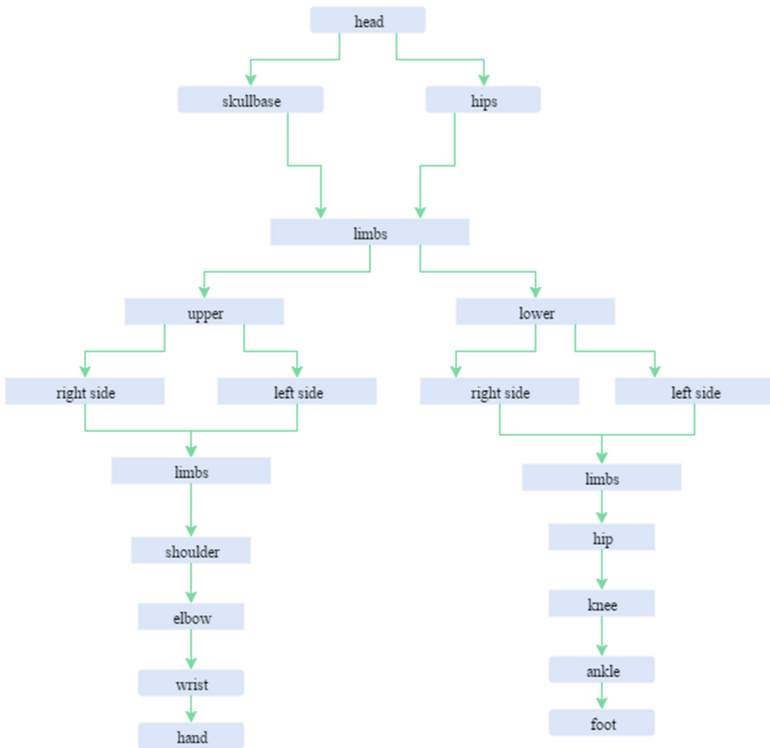


Fig. 4. Elements' hierarchy in our model.

5.1 XML Scheme

The scheme rests in an XML document which holds both the rules for encoding and the way the different elements are related to each other. As an example, consider the XML below, which describes a 30° counter-clockwise rotation of the hand, shaped as a fist, around the axis formed between hand and elbow, measured at the wrist joint. At the same time, the hand rotates clockwise, around the elbow. Within this set-up, even though movements are defined independently, they must be interpreted as one joint movement, thereby establishing relationships between both directly and non-directly connected body parts.

```

1 <node1_angle_between>
2   <hand>
3     <format_attributes>fist</format_attributes>
4     <min_rot_attributs>counter_clockwise</min_rot_attributs>
5     <axis_attributes>x</axis_attributes>
6   </hand>
7   <wrist>
8     <min_rot_attributes>counter_clockwise</min_rot_attributes>
9     <axis_attributes>y</axis_attributes>
10    <angle_attributes>30</angle_attributes>
11  </wrist>
12  <elbow>
13    <min_rot_attributes>clockwise</min_rot_attributes>
14    <axis_attributes>y</axis_attributes>
15  </elbow>
16 </node1_angle_between>
17 . . .
18 <xs: complexType>
19   <xs: sequence>
20     <xs: element name="node1_angle_between">
21       <xs: complexType>
22         <xs: sequence>
23           <xs: element name="hand">
24             <xs: complexType>
25               <xs: sequence>
26                 <xs: group ref="format"/>
27                 <xs: group ref="min_rot"/>
28                 <xs: group ref="axis"/>
29               </xs: sequence>
30             </xs: complexType>
31           </xs: element>
32           <xs: element name="wrist">
33             <xs: complexType>
34               <xs: sequence>
35                 <xs: group ref="min_rot"/>
36                 <xs: group ref="axis"/>
37                 <xs: group ref="angle"/>
38               </xs: sequence>
39             </xs: complexType>
40           </xs: element>

```



```

41     <xs: element name="elbow">
42       <xs: complexType>
43         <xs: sequence>
44           <xs: group ref="min-rot"/>
45           <xs: group ref="axis"/>
46         </xs: sequence>
47       </xs: complexType>
48     </xs: element>
49   </xs: sequence>
50 </xs: complexType>
51 . . .

```

6 Results

In order to help end users in the task of codifying stances and movements without having to dig too deep into the language idiosyncrasies, we developed a computer program which, through a graphical interface, allows them to select body parts and relationships between them from a list of pre-defined values. With this program, the user describes movements by inserting parameters such as axes and angles between joints and by establishing relations between different elements. The program then generates the XML, showing it to the user, who may choose to export it to an XML file. Figures 5 and 6 illustrate the program’s interface. At this point, it is important to notice that the program is still a prototype, available in Portuguese only. New versions are expected soon.

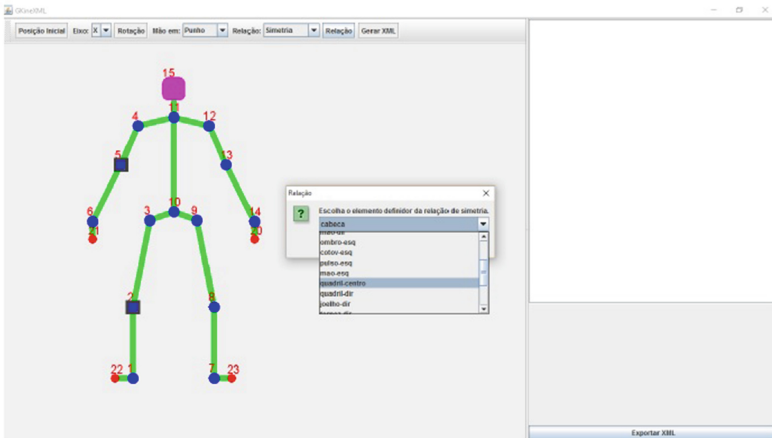


Fig. 5. Program’s interface – defining relationships between elements.

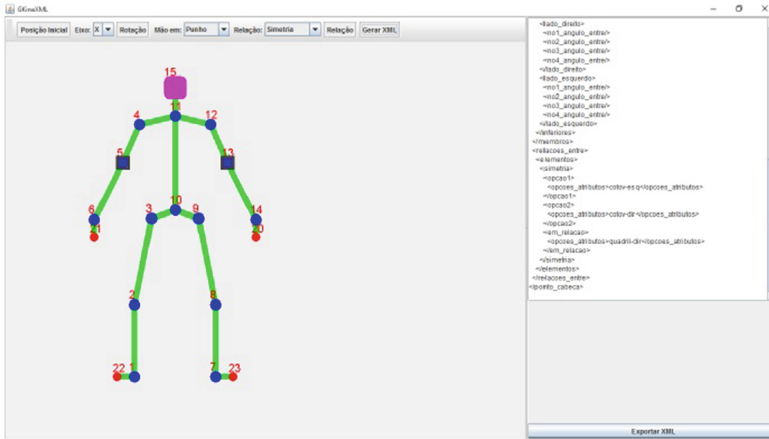


Fig. 6. Program’s interface – outputting the XML code.

To test our prototype we carried out some experiments, where one of the researchers would describe the stances and movements, using the program’s interface, and then manually verify the generated XML file. Coding samples may be found below⁹.

```

1 <head>
2 <limb>
3 <right><!-- angles: upper limbs, right -->
4 <node_angle_between> <!-- angle in the wrist joint-->
5 <hand>
6 <format_attributes> fist</format_attributes>
7 <min_rot_attributes> clockwise </min_rot_attributes>
8 <exis_attributes>x</exis_attributes>
9 </hand>
10 <wrist>
11 <min_rot_attributes> clockwise </min_rot_attributes>
12 <exis_attributes>y</exis_attributes>
13 <angle_attributes>45</angle_attributes>
14 </wrist>
15 <elbow>
16 <min_rot_attributes>clockwise</min_rot_attributes>
17 <exis_attributes>y</exis_attributes>
18 </elbow>
19 </node_angle_between>

```

In this code snippet, we see the description of the right hand rotated clockwise, by 45°, at the right wrist joint, while taking a fist shape. In the following code, the relation of symmetry between left knee and right knee, relatively to the vertical axis passing through the body centre, is illustrated.

⁹ Since generated files are too long to be shown here, we present but snippets for illustration purposes only.

```

1 <relationship_between>
2   <limb>
3     < symmetry >
4       <option1>
5         <options_attributes>left_knee</options_attributes>
6       </option1>
7       <options2>
8         <options_attributes>right_knee</options_attributes>
9       </option2>
10      < regarding>
11        <options_attributes>vertical_axis</options_attributes>
12      </regarding >
13    </symmetry>
14 </relationship _between>

```

7 Conclusion

In this article we proposed a mark-up language for body stance and movement description. Its novelty lies in that it covers many aspects of this subject, from the unambiguous representation of movement to the extraction of relationships between directly and non-directly connected body parts. Along with the language description, we have also developed a computer program to bridge the gap between the end user, that is the person directly dealing with the movements to be codified, and the language details, thereby saving time and reducing the cognitive load imposed by the learning and mastering of the language.

Despite the results we already have, there is still a long road ahead, with much work to be done so as to polish the language, by further testing it and observing its behaviour when dealing with everyday problems in the area. Also, the auxiliary program is still a prototype and, as such, must be tested and further improved, so as to provide the community with a reliable and powerful tool.

References

1. Alexiadis, D.S., Kelly, P., Daras, P., O'Connor, N.E., Boubekeur, T., Moussa, M.B.: Evaluating a dancer's performance using kinect-based skeleton tracking. In: Candan, K.S., Panchanathan, S., Prabhakaran, B., Sundaram, H., Feng, W.C., Sebe, N. (eds.) the 19th ACM International Conference, pp. 659–622 (2011)
2. Bordegoni, M., Camere, S., Caruso, G., Cugini, U.: Body tracking as a generative tool for experience design. In: Duffy, V.G. (ed.) DHM 2015. LNCS, vol. 9185, pp. 122–133. Springer, Heidelberg (2015)
3. Chang, C.Y., Lange, B., Zhang, M., Koenig, S., Requejo, P., Somboon, N., Sawchuk, A., Rizzo, A.: Towards pervasive physical rehabilitation using microsoft kinect. In: Arriaga, R., Matic, A. (eds.) 6th International Conference on Pervasive Computing Technologies for Healthcare (2012)
4. Chang, Y.J., Chen, S.F., Huang, J.D.: A kinect-based system for physical rehabilitation: a pilot study for young adults with motor disabilities. *Res. Dev. Disabil.* **32**(6), 2566–2570 (2011)

5. Chye, C., Sakamoto, M., Nakajima, T.: An exergame for encouraging martial arts. In: Kurosu, M. (ed.) *HCI 2014, Part III. LNCS*, vol. 8512, pp. 221–232. Springer, Heidelberg (2014)
6. Clark, R.A., Pua, Y.H., Fortin, K., Ritchie, C., Webster, K.E., Denehy, L., Bryant, A.L.: Validity of the microsoft kinect for assessment of postural control. *Gait Posture* **36**(3), 372–377 (2012)
7. Eshkol, N., Harries, J.G., Sapir, T., Sella, R., Shoshani, M.: *The Quest for Tai Chi Chuan*. The Research Center for Movement Notation, Faculty of Visualand Performing Arts, Tel Aviv University (1986)
8. Guest, A.H.: *Labanotation: The System of Analyzing and Recording Movement*. Taylor & Francis, London (2014)
9. Hanke, T.: Hamnosys - representing sign language data in language resources and language processing contexts. In: *Proceedings of the 4th International Conference on Language Resources and Evaluation (LREC 2004)*. Lisbon, Portugal, 26–28 May 2004
10. Huang, Z., Eliëns, A., Visser, C.: Step: a scripting language for embodied agents. In: *Proceedings of the Workshop of Lifelike Animated Agents*, Tokyo (2002)
11. Huang, Z., Eliëns, A., Visser, C.: XSTEP: An XML-based markup language for embodied agents. In: *Proceedings of the 16th International Conference on Computer Animation and Social Agents, (CASA 2003)*, pp. 105–110 (2003)
12. Huang, Z., Eliëns, A., Visser, C.: Step: a scripting language for embodied agents. In: Gabbay, M., Siekmann, J., Prendinger, H., Ishizuka, M. (eds.) *Life-Like Characters. Cognitive Technologies*, pp. 87–109. Springer, Heidelberg (2004)
13. Kennaway, R.: Avatar-independent scripting for real-time gesture animation. arXiv preprint (2015). [arXiv:1502.02961](https://arxiv.org/abs/1502.02961)
14. Kleinman, S.: Movement notation systems: An introduction. *Quest* **23**(1), 33–34 (1975)
15. Rett, J., Dias, J., Ahuactzin, J.M.: Laban movement analysis using a bayesian model and perspective projections. In: *Proceedings of the Brain, Vision and AI* (2008)
16. Trajkova, M., Ferati, M.: Usability evaluation of kinect-based system for ballet movements. In: Marcus, A. (ed.) *DUXU 2015. LNCS*, vol. 9187, pp. 464–472. Springer, Heidelberg (2015)