UNIVERSIDAD DE OVIEDO

PROGRAMA DE DOCTORADO EN INGENIERÍA INFORMÁTICA

CONTRIBUCIÓN TEÓRICO-PRÁCTICA EN EL MODELADO LINGÜÍSTICO DE FENÓMENOS COMPLEJOS

Daniel Sánchez Valdés

UNIVERSIDAD DE OVIEDO

PROGRAMA DE DOCTORADO EN
INGENIERÍA INFORMÁTICA

CONTRIBUCIÓN TEÓRICO-PRÁCTICA EN EL MODELADO LINGÜÍSTICO DE FENÓMENOS COMPLEJOS

Daniel Sánchez Valdés

Director:
Gracián Triviño Barros

Julio 2015

RESUMEN DEL CONTENIDO DE TESIS DOCTORAL

1.- Título de la Tesis	
Español: Contribución teórico-práctica en el modelado lingüístico de fenómenos complejos	Inglés: Theoretical and practical contribution to the linguistic modeling of complex phenomena

2.- Autor	
Nombre: Daniel Sánchez Valdés	DNI/Pasaporte/NIE:
Programa de Doctorado: Ingeniería Informática	
Órgano responsable: Comisión Académica del Programa de Doctorado	

RESUMEN (en español)

Las nuevas tecnologías permiten adquirir y almacenar grandes volúmenes de datos de distinta naturaleza. Con el objetivo de transformar estos datos en información, deben ser analizados y explicados de la manera más comprensible e interpretable posible, empleando para ello todo el conocimiento que se posea sobre el fenómeno estudiado. Gracias al lenguaje natural podemos construir sistemas computacionales que interpreten y describan las características más relevantes de un fenómeno imitando la forma que tienen los seres humanos de percibir e interpretar la información. Esta descripción se realiza recopilando las expresiones lingüísticas que se emplean comúnmente para definir el fenómeno, y escogiendo en cada momento la expresión o expresiones que mejor definan la situación, esto es, que presenten el mayor grado de validez.

El objetivo principal de esta tesis es extender la Teoría Computacional de Percepciones, introducida por Lotfi Zadeh en 1999. La idea consiste en utilizar la Lógica Difusa para crear modelos basados en la forma que tienen los seres humanos de hacer descripciones a través del lenguaje natural. El objetivo es utilizar las estructuras del lenguaje para hacer modelos precisos y robustos que manejen la imprecisión inherente de los fenómenos complejos. Durante los últimos años, la unidad de Computación con Percepciones del *European Centre for Soft Computing* ha desarrollado sistemas computaciones que utilizan el lenguaje natural para describir lingüísticamente los datos recogidos de un fenómeno. Su principal contribución en este campo de investigación consiste en la creación del denominado *Granular Linguistic Model of a Phenomenon*, que analiza los datos de entrada, combina diferentes fuentes de conocimiento y genera informes sobre los aspectos más relevantes del fenómeno estudiado, adaptando el contenido al receptor final del mensaje.

Esta tesis recoge como testigo los primeros resultados obtenidos en esta dirección, incorporando mejoras y nuevos elementos en la arquitectura general de modelado lingüístico. La metodología llevada a cabo ha consistido en el desarrollo e incorporación de nuevos conceptos teóricos, como son el "desconocimiento" y la "relevancia", dos características intrínsecas en el modelado de fenómenos complejos, que hasta la actualidad no se habían tenido en cuenta. Además, se ha ampliado y evolucionado el concepto de *Fuzzy Finite State Machine*, un paradigma para modelar aquellos fenómenos que evolucionan en el tiempo, generalmente de manera quasi-periódica.

Por otra parte, el empleo de emociones en el desarrollo de sistemas computacionales juega un papel fundamental en los procesos de toma de decisiones. La sensibilidad emocional y la expresividad son muy importantes en la interacción humana y transmiten información valiosa. Dado que la descripción lingüística de datos está especialmente pensada, en general, para aplicaciones en las que existe una fuerte interacción hombre-máquina, esta tesis se ha adentrado en el mundo del *Affective Computing*, modelando emocionalmente un agente virtual cuyo carácter (gestos faciales y expresiones lingüísticas) evolucionan de acuerdo al grado con el que se alcanza el objetivo perseguido.

Cada contribución teórica se ha respaldado con la resolución de problemas prácticos con aplicación en el mundo real. De este modo, se ha creado un modelo de la marcha humana más completo y robusto, diseñando y elaborando una aplicación capaz de describir la calidad de nuestra forma de caminar. Asimismo, se ha creado un modelo que analiza las principales actividades del ser humano, con el objetivo de proporcionar una herramienta de auto-seguimiento de la actividad diaria. Finalmente, se han aplicado estas técnicas para modelar fenómenos de otros campos como la astronomía, el transporte o el consumo eléctrico.

RESUMEN (en Inglés)

New technologies allow us to acquire and store big amounts of data from very different fields. With the aim of transform these data into information, they must be analyzed and explained in an interpretable and understandable way, using for this purpose the available knowledge about the phenomenon. Thanks to the natural language, we can build computational systems that interpret and describe the more relevant characteristics of the phenomenon, in the same way human beings perceive and interpret the information. This description is done by taking the most commonly expressions used to define the phenomenon, choosing the expression or expressions that better define the situation, i.e., those which have the highest validity degree.

The main goal of this Ph.D. thesis is to extend the Computational Theory of Perceptions, introduced by Lotfi Zadeh in 1999. The idea consists in using Fuzzy Logic to create models based in the way human beings make descriptions by means of natural language. The goal is to use the language structures to design precise and robust models that deals with the imprecision inherent in complex phenomena. During the last years, the research unit “Computing with Perceptions” of the European Centre for Soft Computing, has developed computational systems that use natural language to linguistically describe the data acquired from a phenomenon. Its main contribution in this field is the so called “Granular Linguistic Model of a Phenomenon”, that analyzes the input data, combines different knowledge sources and generates reports about the relevant aspects of the studied phenomenon, adapting the content to the final receptor of the message.

This Ph.D. thesis takes the first results obtained in this direction, providing new advances and new elements to the general architecture of linguistic modeling. Methodology followed has consisted in developing and incorporating new theoretical concepts, as the “ignorance” and the “relevance”. In addition, the paradigm of Fuzzy Finite State Machine has been improved, in order to model those phenomena that evolve in time, generally in a quasi-periodic way.

By other hand, the use of emotions in the development of computational systems plays a fundamental role in decision-making processes. The emotional sensibility and expressivity are very important in the human interaction and transmit valuable information. Since the linguistic description of data is specially thought, in general, for those applications with a strong human-computer interaction, this Ph.D. thesis has deepened in the field of Affective Computing, emotionally modeling a virtual agent which mood (facial and linguistic expressions) evolve according to the objectives fulfillment.

Each theoretical contribution has been supported by the resolution of practical problems with real demand. In this way, a more complete and robust model of human gait has been developed, designing an application able to describe the gait quality. Also, a new model has been created to analyze the main activities of humans, with the aim of providing a self-tracking tool for physical activity. Finally, these techniques have been applied to model phenomena of other fields, such as astronomy, transport or electricity consumption.

SR. DIRECTOR DE DEPARTAMENTO DE INFORMÁTICA
SR. PRESIDENTE DE LA COMISIÓN ACADÉMICA DEL PROGRAMA DE DOCTORADO EN INGENIERÍA INFORMÁTICA

Agradecimientos

Esta tesis doctoral no habría sido posible sin la ayuda y el apoyo de muchas personas. A todas y cada una de ellas les quiero agradecer de corazón que me hayan ayudado a lo largo del camino.

En particular, quiero agradecer a mi director Gracián Triviño por ofrecerme la posibilidad de realizar el doctorado bajo su dirección, por todos sus consejos y enseñanzas, tanto a nivel científico como a nivel personal, así como por su apoyo y dedicación durante todos estos años.

De igual modo, quiero agradecer al European Centre for Soft Computing haberme ofrecido la oportunidad de realizar esta tesis doctoral bajo su cobijo, así como a todos los compañeros que día a día me han respaldado y acompañado, contribuyendo sin lugar a dudas al desarrollo de este trabajo.

Finalmente, agradecer a mi familia todo su apoyo y comprensión durante estos años de investigación. Con ellos he compartido todas las penas y alegrías del camino. Parte de este trabajo es suyo, pues han sido mis ojos, mis oídos y mi aliento en los momentos en los que más lo necesitaba.

Es necesario resaltar que gran parte del trabajo presentado en esta tesis doctoral ha sido financiado por el proyecto del Ministerio de Ciencia e Innovación bajo el proyecto "Descripción Lingüística de Fenómenos Complejos" (TIN2011-29827-C02-01) y por la Consejería de Economía y Empleo del Principado de Asturias bajo el proyecto "Descripción Lingüística de Fenómenos Complejos" (COF13-054).

Nunca consideres el estudio como una obligación, sino como una oportunidad para penetrar en el bello y maravilloso mundo del saber.
Albert Einstein (1879 - 1955)

Índice general

Resumen

Las nuevas tecnologías permiten adquirir y almacenar grandes volúmenes de datos de distinta naturaleza. Con el objetivo de transformar estos datos en información, deben ser analizados y explicados de la manera más comprensible e interpretable posible, empleando para ello todo el conocimiento que se posea sobre el fenómeno estudiado.

Gracias al lenguaje natural podemos construir sistemas computacionales que interpreten y describan las características más relevantes de un fenómeno imitando la forma que tienen los seres humanos de percibir e interpretar la información. Esta descripción se realiza recopilando las expresiones lingüísticas que se emplean comúnmente para definir el fenómeno, y escogiendo en cada momento la expresión o expresiones que mejor definan la situación, esto es, que presenten el mayor grado de validez.

El objetivo principal de esta tesis es extender la Teoría Computacional de Percepciones, introducida por Lotfi Zadeh en 1999. La idea consiste en utilizar la Lógica Difusa para crear modelos basados en la forma que tienen los seres humanos de hacer descripciones a través del lenguaje natural. El objetivo es utilizar las estructuras del lenguaje para hacer modelos precisos y robustos que manejen la imprecisión inherente de los fenómenos complejos. Durante los últimos años, la unidad de Computación con Percepciones del *European Centre for Soft Computing* ha desarrollado sistemas computaciones que utilizan el lenguaje natural para describir lingüísticamente los datos recogidos de un fenómeno. Su principal contribución en este campo de investigación consiste en la creación del denominado *Granular Linguistic Model of a Phenomenon* (Modelo Lingüístico Granular de un Fenómeno), que analiza los datos de entrada, combina diferentes fuentes de conocimiento y genera informes sobre los aspectos más relevantes del fenómeno estudiado, adaptando el contenido al receptor final del mensaje.

Esta tesis recoge como testigo los primeros resultados obtenidos en esta dirección, incorporando mejoras y nuevos elementos en la arquitectura general de modelado lingüístico. La metodología llevada a cabo ha consistido en el desarrollo e incorporación de nuevos conceptos teóricos, como son el "desconocimiento" y la "relevancia", dos características intrínsecas en el modelado de fenómenos complejos, que hasta la actualidad no se habían tenido en cuenta. Además, se ha ampliado y evolucionado el concepto de *Fuzzy Finite State Machine* (Máquina de Estados Difusos Finitos), un paradigma para modelar aquellos fenómenos que evolucionan en el tiempo, generalmente de manera quasi-periódica.

Por otra parte, el empleo de emociones en el desarrollo de sistemas computacionales juega un papel fundamental en los procesos de toma de decisiones. La sensibilidad emocional y la expresividad son muy importantes en la interacción humana y transmiten información valiosa. Dado que la descripción lingüística de datos está especialmente pensada, en general, para aplicaciones en las que existe una fuerte interacción hombre-máquina, esta tesis se ha adentrado en el mundo del *Affective Computing*, modelando emocionalmente un agente virtual cuyo carácter (gestos faciales y expresiones lingüísticas) evolucionan de acuerdo al grado con el que se alcanza el objetivo perseguido.

Cada contribución teórica se ha respaldado con la resolución de problemas prácticos con aplicación en el mundo real. De este modo, se ha creado un modelo de la marcha humana más completo y robusto, diseñando y elaborando una aplicación capaz de describir la calidad de nuestra forma de caminar. Asimismo, se ha creado un modelo que analiza las principales actividades del ser humano, con el objetivo de proporcionar una herramienta de auto-seguimiento de la actividad diaria. Finalmente, se han aplicado estas técnicas para modelar fenómenos de otros campos como la astronomía, el transporte o el consumo eléctrico.

Parte I

Memoria

Capítulo 1

Introducción

Entendemos por "fenómeno" cualquier manifestación física que pueda contrastarse por observación directa o medición indirecta (fenómenos geológicos, eléctricos, químicos, biológicos, ...), así como cualquier situación constatable mediante la recopilación de informes o encuestas (fenómenos psicológicos, sociológicos o económicos).

Las nuevas tecnologías permiten adquirir y almacenar grandes volúmenes de datos de diferente naturaleza. Con el objetivo de transformar estos datos en información, deben ser analizados y explicados de la manera más comprensible e interpretable posible, empleando para ello todo el conocimiento que se posea sobre el fenómeno estudiado. Sin embargo, la relación entre la cantidad de información disponible y el número de expertos capaces de analizar dicha información crece dramáticamente, siendo de vital importancia desarrollar sistemas computacionales capaces de interpretar y describir los datos más relevantes.

La tarea de modelar fenómenos complejos se conoce como Identificación de Sistemas (IS) [Sod 94, Lju 98]. Tanto el concepto de IS como su definición formal fueron introducidos por Lotfi Zadeh en 1956 [Zad 56]. De acuerdo a la definición que dio Zadeh [Zad 62], dada una clase de modelos, la IS requiere encontrar un modelo que pueda considerarse equivalente al sistema objetivo con respecto a los datos de entrada/salida. La IS emplea métodos estadísticos para construir modelos matemáticos de sistemas dinámicos a partir de los datos medidos. Un modelo matemático dinámico en este contexto es una descripción matemática del comportamiento dinámico de un sistema o proceso, tanto en el dominio de la frecuencia como en el del tiempo.

Tradicionalmente en IS, los ingenieros emplean ecuaciones diferenciales para construir modelos de "caja blanca" basados en el principio de modelar el comportamiento de sistemas del mundo real [Oga 67, Lju 98, Nel 00, Ise 09]. Desafortunadamente, la mayoría de los sistemas de nuestro entorno son complejos y no es posible, o es muy costoso, obtener y resolver estas ecuaciones. Por lo tanto, cuando el sistema modelado crece en complejidad, el número de variables y ecuaciones necesarias se vuelve inmanejable y no existe otra opción que trabajar con modelos alternativos. Esta situación la describe Zadeh en su Principio de Incompatibilidad: *"cuanto más se incremente la complejidad de un sistema, nuestra habilidad para hacer declaraciones precisas sobre su comportamiento disminuye* *hasta que se alcanza un umbral más allá del cual la precisión y la importancia (o relevancia) se vuelven casi características mutuamente excluyentes"* [Zad 73].

Una alternativa posible consiste en analizar el comportamiento del sistema y sus influencias externas (entradas del sistema) y tratar de determinar una relación matemática entre ellas sin tener en cuenta su funcionamiento interno. Estos modelos se conocen como modelos de "caja negra". En otras palabras, en un modelo de "caja negra" nos interesa su forma de interactuar con el entorno que lo rodea entendiendo qué es lo que hace pero no cómo lo hace. Algunos ejemplos de este tipo de modelos son las redes neuronales [Hay 94] o los modelos lineales autoregresivos [Sea 97, Sta 09]. Debido a que existen gran cantidad de aplicaciones que requieren un conocimiento exhaustivo del modelo, este tipo de soluciones generan cierta desconfiaza y se buscan alternativas intermedias, como son los modelos de "caja gris". En estos modelos se combinan algunos detalles sobre el comportamiento interno del sistema con la experiencia y el conocimiento extraídos de los datos de entrada.

En 1965, Zadeh introdujo el concepto de conjunto difuso y, algunos años más tarde, estableció lo que hoy conocemos como Lógica Difusa o *Fuzzy Logic*, que nació con el objetivo de manejar la imprecisión y la variabilidad inherente a los datos que manejamos, así como modelar aquellos sistemas en los cuales obtener y resolver apropiadamente las ecuaciones de estado resulta una tarea difícil o imposible de realizar. Pero en algunos sistemas, el número de variables y de reglas necesarias para crear modelos difusos crecía hasta volverse prácticamente incomprensibles y, en consecuencia, difícilmente aplicables. Para solucionar este problema, Zadeh propuso el uso de variables lingüísticas [Zad 75] en vez de, o además de, valores numéricos, así como la caracterización de las relaciones entre las variables a través de declaraciones difusas condicionales.

El modelado difuso es uno de los asuntos más importantes dentro de la Lógica Difusa y se utiliza para describir el comportamiento de un sistema mediante el uso del lenguaje natural. El empleo de variables lingüísticas y reglas en lenguaje natural [Mam 74, Mam 75, Mam 77] reduce notablemente el esfuerzo existente a la hora de extraer conocimiento experto. Por otra parte, siendo aproximadores universales [Buc 93, Cas 95], los sistemas de inferencia difusa son capaces de realizar mapeos lineales entre las entradas y las salidas.

Gracias al lenguaje natural podemos construir sistemas computacionales que interpreten y describan las características más relevantes de un fenómeno imitando la forma que tienen los seres humanos de percibir e interpretar la información. La descripción lingüística de datos está especialmente diseñada para aquellas aplicaciones en las que existe una fuerte interacción entre las personas y las máquinas, cuando deben interpretarse grandes volúmenes de datos, a priori, difíciles de comprender.

Aunque el concepto básico de "resúmenes lingüísticos difusos" se estableció en 1982 por Ronald Yager, fue Zadeh quien en 1999 estableció la Teoría Computacional de Percepciones con su artículo *"From computing with numbers to computing with words - from manipulation of measurements to manipulation of perceptions"*, ampliada en sus artículos posteriores. De acuerdo con la Gramática Sistémica Funcional, el lenguaje natural es un sistema que se organiza en cuatro capas o estratos, conocidos como Contexto, Semántica, Léxico-gramática y Expresión. Los trabajos más recientes sobre el procesamiento del lenguaje natural, incluyendo aquellos basados en Lógica Difusa, se centran en la capa Léxico-gramatical. El enfoque de la unidad de Computación con Percepciones del *European Centre for Soft Computing* (ECSC) se centra especialmente en el modelado del Contexto y la Semántica. El diseñador debe realizar un estudio sobre el uso que se hace del lenguaje natural en el contexto específico de cada aplicación, de lo cual se extraerá un *corpus* con todas las posibles expresiones utilizadas comúnmente para describir el fenómeno monitorizado. Este *corpus* de expresiones lingüísticas cubre el uso del lenguaje natural en todo el conjunto posible de situaciones (Contexto) y en el conjunto completo de percepciones relevantes (Semántica).

A lo largo de los años, los modelos basados en Lógica Difusa han crecido en complejidad como consecuencia de los requisitos de modelado en términos de precisión e interpretabilidad. Actualmente existen numerosas investigaciones en esta dirección, trabajando por establecer un formalismo que permita que los modelos difusos diseñados sean lo más comprensibles por el ser humano [Cas 03b, Cas 03a, Alo 07, Alo 08, Alo 11a, Alo 11b].

1.1. Punto de partida de la tesis

Esta tesis continúa la línea investigadora desarrollada por Alberto Álvarez Álvarez en su tesis doctoral titulada *"Linguistic modeling of complex phenomena"*. En esta sección se presentan brevemente los elementos que componen la arquitectura general del modelado lingüístico, con el objetivo de que el lector comprenda mejor las contribuciones, tanto teóricas como prácticas, desarrolladas.

De acuerdo con Zadeh, el objeto de las percepciones no son solo los atributos de los objetos, como la distancia, la velocidad o el ángulo, sino que pueden ser los sistemas completos, por ejemplo, una persona aparcando su coche o el tráfico en una rotonda, lo cual es percibido por un entorno informatizado. Los fenómenos pertenencen a un determinado contexto y evolucionan a lo largo del tiempo entre los distintos tipos de situaciones.

El Modelo Lingüístico Granular de un Fenómeno se basa en las percepciones subjetivas de un dominio experto que podemos denominar diseñador. Cuanta más experiencia posea el diseñador, mejor comprensión y mayor uso del lenguaje natural, más rico será el modelo y más posibilidades de conseguir satisfacer las necesidades del usuario final, así como de responder a sus expectativas. Para ello, el diseñador emplea los recursos disponibles, por ejemplo los sensores, para adquirir datos sobre el fenómeno monitorizado, y utiliza su propia experiencia para interpretar estos datos y crear el modelo. Luego, utiliza los recursos computacionales para producir las expresiones lingüísticas que describan lo que está ocurriendo en cada instante de tiempo. En las subsecciones siguientes, se introducen los principales elementos de la arquitectura desarrollada para describir lingüísticamente fenómenos complejos.

1.1.1. Computational Perception

Una *Computational Perception* (CP) es el modelo computacional de una unidad de información adquirida por el diseñador sobre el fenómeno que quiere modelar. Está basada en el concepto de variable lingüística de Zadeh [Zad 75]. En general, las CPs recogen aspectos particulares del fenómeno con distintos niveles de granularidad. Una CP está formada por la tupla (A, W), cuyos componentes se describen de la siguiente manera:

$A = (a_1, a_2, \ldots, a_n)$ es un vector de expresiones lingüísticas (palabras y sentencias en lenguaje natural) que representan el dominio lingüístico completo de una CP. Los componentes de A están definidos por el diseñador de acuerdo con las expresiones más apropiadas de entre las utilizadas típicamente en el dominio del lenguaje de la aplicación. De este modo, cada componente a_i corresponde con el valor lingüístico de CP más adecuado para cada situación con una granularidad específica. Cuanto mayor sea el número de expresiones, mayor serán la granularidad y la precisión obtenidas. Estas expresiones pueden ser tan simples como, por ejemplo, $a_i =$ *"La temperatura es bastante alta"*, o complejas, $a_i =$ *"Hoy el tiempo está mejor que durante los últimos días"*. Las expresiones lingüísticas a_i se utilizan para generar informes lingüísticos sobre el fenómeno monitorizado.

$W = (w_1, w_2, \ldots, w_n)$ es un vector de grados de validez $w_i \in [0, 1]$ asignados a cada a_i. En el contexto de aplicación, w_i representa la idoneidad de a_i para describir la percepción actual. El diseñador escoge los componentes de A para tratar de cubrir todos los posibles valores del CP, es decir, la validez total debe ser distribuida entre todas las etiquetas lingüísticas. Por lo tanto, típicamente, los componentes de A están sujetos a conjuntos difusos que forman particiones difusas fuertes [Rus 69] en el universo del discurso del CP, es decir, $\sum_{i=1}^{n} w_i = 1$.

1.1.2. Perception Mapping

Un *Perception Mapping* (PM) se utiliza para combinar o agregar CPs. Está formado por la tupla (U, y, g, T), en la cual:

U es un vector de valores numéricos o CPs de entrada. En el primer caso, las entradas al modelo $U = (z_1, z_2, \ldots, z_m)$ son valores numéricos $z_i \in \mathbb{R}$ adquiridos mediante sensores o extraídos de una base de datos, y se les conoce como PMs de primer orden (1-PMs). En el segundo caso, las entradas al modelo $U = (u_1, u_2, \ldots, u_m)$ son CPs, donde u_i son tuplas (A_i, W_i) y m representa el número de CPs de entrada. A este tipo de PMs se les conoce como PMs de segundo orden (2-PMs).

$y = (A_y, W_y)$ es la CP de salida. De este modo, PM: $u_1 \times u_2 \times \cdots \times u_m \longrightarrow y$. Llamamos CPs de primer orden (1-CPs) a las salidas de las 1-PMs, y CPs de segundo orden (2-CPs) a las salidas de las 2-PMs.

g es una función de agregación $W_y = g(W_1, W_2, \ldots, W_m)$, donde W_i son los vectores de los grados de validez de las m CPs de entrada.

En Lógica Difusa se han desarrollado muchos tipos distintos de funciones de agregación. Por ejemplo, g puede ser implementado usando un conjunto de reglas difusas. En el caso de 1-PMs, g se construye usando un conjunto de funciones de pertenencia $\mu_{a_i}(z)$ de la siguiente manera: $W_y = (\mu_{a_1}(z), \mu_{a_2}(z), \ldots, \mu_{a_n}(z)) = (w_1, w_2, \ldots, w_n)$

T es un algoritmo de generación de texto que permite generar las sentencias recogidas en A_y. T es habitualmente una plantilla lingüística sencilla, como por ejemplo, *"La temperatura de esta habitación es* $\{$*baja* | *media* | *alta*$\}$*"*.

1.1.3. Granular Linguistic Model of a Phenomenon

Un *Granular Linguistic Model of a Phenomenon* (GLMP) es una red de PMs encargados de combinar o agregar CPs. Se define como un paradigma utilizado para desarrollar sistemas computacionales capaces de generar descripciones lingüísticas a partir de los datos de entrada.

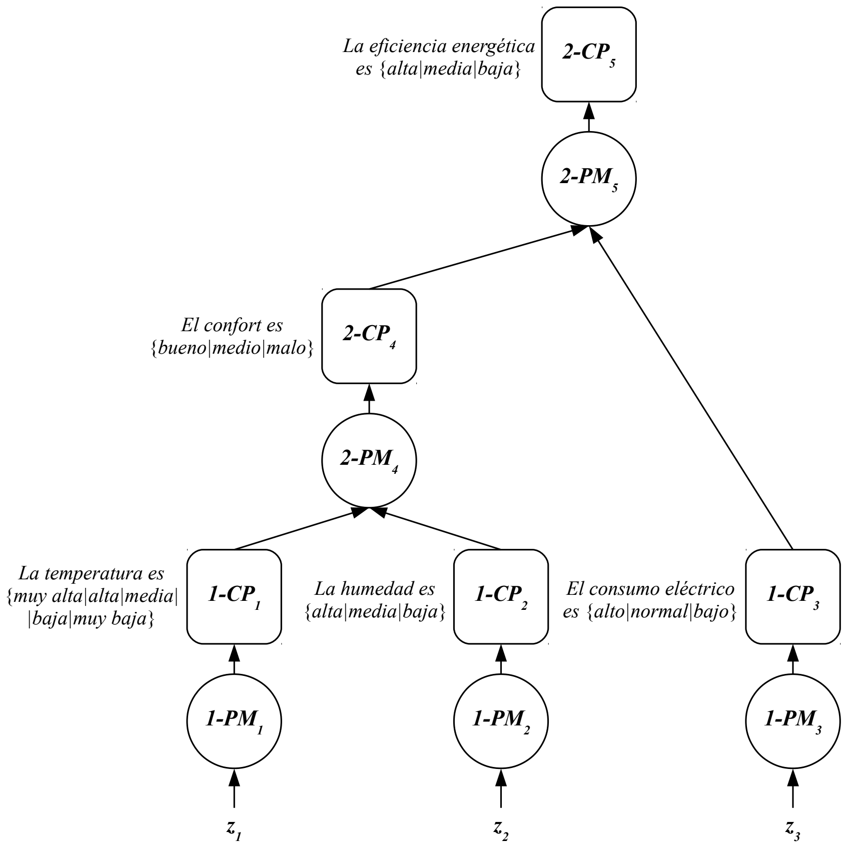


Figura 1.1: Ejemplo de un GLMP sencillo que modela la eficiencia energética.

A raíz de utilizar distintas funciones de agregación y distintas expresiones lingüísticas, el paradigma del GLMP permite al diseñador modelar computacionalmente sus percepciones. Es importante destacar que, después de ser instanciado con un conjunto de datos de entrada, el GLMP proporciona una estructura que, en aplicaciones de tamaño medio, podría incluir cientos de sentencias válidas.

La Fig. 1.1 muestra un sencillo GLMP que modela la eficiencia energética de la climatización de una habitación, en base a percepciones como el confort, la temperatura, la humedad y el consumo eléctrico. En este ejemplo, las variables de entrada proporcionan en cada instante de tiempo los valores z_1, z_2, y z_3. Estas variables se introducen en 1-PM_1, 1-PM_2 y 1-PM_3, proporcionando 1-CP_1, 1-CP_2 y 1-CP_3. Basándonos en 1-CP_1 y 1-CP_2 usamos 2-PM_4 para explicar 2-CP_4. Finalmente, se proporciona una descripción del fenómeno en el nivel más alto de abstracción, gracias a 2-CP_5, la cual es explicada por 2-PM_5 a partir de 1-CP_3 y 2-CP_4. Notar que, utilizando esta estructura no solo se puede proporcionar una descripción lingüística del fenómeno a cierto nivel, sino una explicación en terminos lingüísticos al nivel más bajo.

1.1.4. Fuzzy Finite State Machine

En Identificación de Sistemas los diseñadores deciden qué paradigma utilizan para representar los modelos del sistema. La representación de espacios de estados es una de las estructuras de modelos más expresivas, donde el diseñador utiliza su creatividad y su experiencia personal para escoger el conjunto de variables de estado suficientes y necesarias que representen el sistema en cada instante de tiempo [Oga 67].

Cuando el sistema evoluciona en el tiempo, el estado actual sigue una trayectoria en el espacio de estado. Dentro de este contexto, una *Fuzzy Finite State Machine* (FFSM) puede ser considerada como un caso particular de un modelo de espacio de estado, que es especialmente útil cuando se modelan fenómenos que evolucionan en el tiempo mediante un número limitado de estados difusos. El concepto inicial de FFSM fue introducido por Santos [San 68] y posteriormente desarrollado por diferentes autores (por ejemplo, [Mor 02]). Esta familia de FFSMs se caracterizaba por tener estados difusos pero entradas *crisps*. Más tarde, el modelo inicial se extendió para introducir entradas difusas [Yin 02], [Cao 07]. Sin embargo, el concepto de FFSM desarrollado en esta línea de investigación se inspira en los conceptos de estado difuso y sistema difuso desarrollados por Zadeh [Zad 62], [Zad 96]. Más concretamente, se puede considerar como una implementación de la idea general de "modelos difusos entrada-salida de sistemas dinámicos" propuesta por Yager [Yag 94].

Las FFSMs presentadas se utilizan en el modelado de sistemas para permitir a los expertos construir modelos lingüísticos difusos comprensibles de una manera sencilla. Una FFSM es una tupla $\{Q, U, f, Y, g\}$, donde:

Q es el vector de estados del sistema, que se define como una variable lingüística que toma valores dentro del conjunto de etiquetas $(q_1, q_2, \ldots, q_n)$, siendo n el número de estados. Cada uno de los estados representa el patrón de una situación repetitiva dentro del sistema.

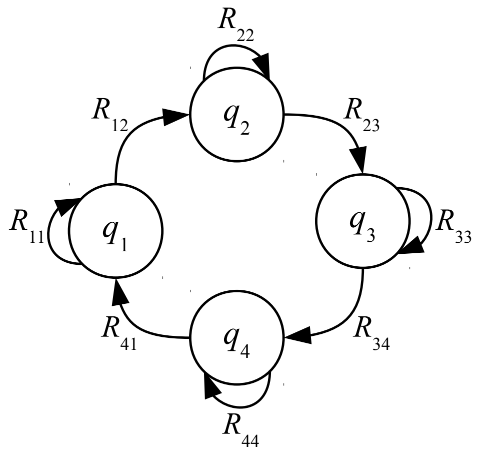


Figura 1.2: Ejemplo de un diagrama de estados.

U es el vector de entrada del sistema $(u_1, u_2, ..., u_{n_u})$, siendo n_u el número de variables de entrada. Estas variables de entrada pueden ser directamente valores numéricos proporcionados por sensores o extraídos de una base de datos, así como CPs provenientes de la agregación o combinación de información en las capas inferiores de modelado. Esta situación es habitual cuando la FFSM actúa como la función de agregación g de un PM dentro de un GLMP.

f es la función de transición que calcula en cada instante t el valor que tiene cada estado perteneciente a Q. Esta función es implementada a partir de una base de conocimiento difusa. Una vez que el experto ha identificado cuál es el vector de estados Q, se deben definir las reglas difusas que gobiernan la evolución temporal del sistema por los estados.

Podemos encontrar reglas para permanecer en un estado q_i (R_{ii}) y reglas para cambiar de un estado q_i a un estado q_j (R_{ij}). El diagrama de estados es el encargado de establecer qué transiciones están permitidas y, por lo tanto, tendrán asociada una regla de transición. Si una transición no figura en el diagrama de estados, no tendrá ninguna regla asociada. La Fig. 1.2 muestra como ejemplo el diagrama de estados de un sistema con cuatro posibles estados de salida.

La expresión genérica de una regla R_{ij} es la siguiente:

$$R_{ij}: \text{ IF } (State[t] \text{ is } q_i) \text{ AND } C_{ij} \text{ THEN } (State[t+1] \text{ is } q_j) \tag{1.1}$$

donde:

1. El primer término del antecedente ($State[t]$ is q_i) representa el grado de activación del estado q_i en el instante de tiempo t.
2. El segundo término del antecedente (C_{ij}) describe las condiciones impuestas sobre las variables de entrada para que se produzca un cambio del estado q_i al

estado q_j.

3. Finalmente, el consecuente ($State[t+1]$ is q_j) representa el grado de activación del estado q_j asociado a esta regla en el instante de tiempo $t+1$.

Y es el vector de salida del sistema $(y_1, y_2, ..., y_{n_y})$, siendo n_y el número de variables de salida.

g es la función de salida que calcula el vector de salida del sistema.

1.1.5. Arquitectura general

La Fig. 1.3 muestra la arquitectura básica necesaria para generar informes. Los principales módulos de procesamiento de este sistema computacional son el módulo de adquisición de datos, el módulo de validación y el módulo de expresión, los cuales se describen brevemente a continuación:

- Módulo de adquisición de datos: este módulo de procesamineto proporciona los datos necesarios para alimentar las 1-CPs. Además, proporciona el interfaz con el entorno físico de la aplicación. Este módulo puede emplear sensores o bien alimentarse de la información contenida en una base de datos.
- Módulo de validación: una vez que los datos de entrada están disponibles, el módulo de validación usa las funciones de agregación del GLMP para calcular los grados de validez de cada CP. Por lo tanto, este módulo proporciona como salida una colección de expresiones lingüísticas con sus respectivos grados de validez asociados.
- Módulo de expresión: proporcionadas un conjunto de expresiones lingüísticas, el objetivo es combinar esta información para construir un informe lingüístico. Este módulo se encarga de generar el informe lingüístico más relevante, escogiendo las expresiones más adecuadas y basándose en la estructura fijada por el *Report Template*.

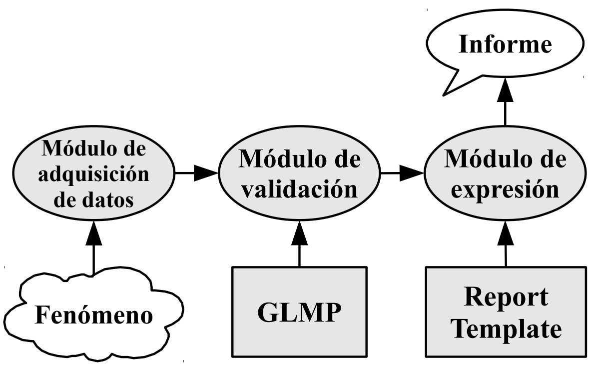


Figura 1.3: Componentes principales de la arquitectura general.

1.2. Trayectoria y contribución

La Fig. 1.4 representa gráficamente la influencia que han tenido las diferentes disciplinas de la Ciencia Cognitiva en el diseño de la arquitectura general para la descripción lingüística de fenómenos complejos, presentada en la sección anterior. En letra azul están reflejados algunos de los trabajos más relevantes de cada una de estas disciplinas; en verde las principales contribuciones realizadas durante los últimos años desde la unidad de Computación con Percepciones del ECSC; y finalmente, en rojo, las contribuciones desarrolladas durante la presente tesis doctoral, presentadas algunas de ellas en la Parte II. Publicaciones.

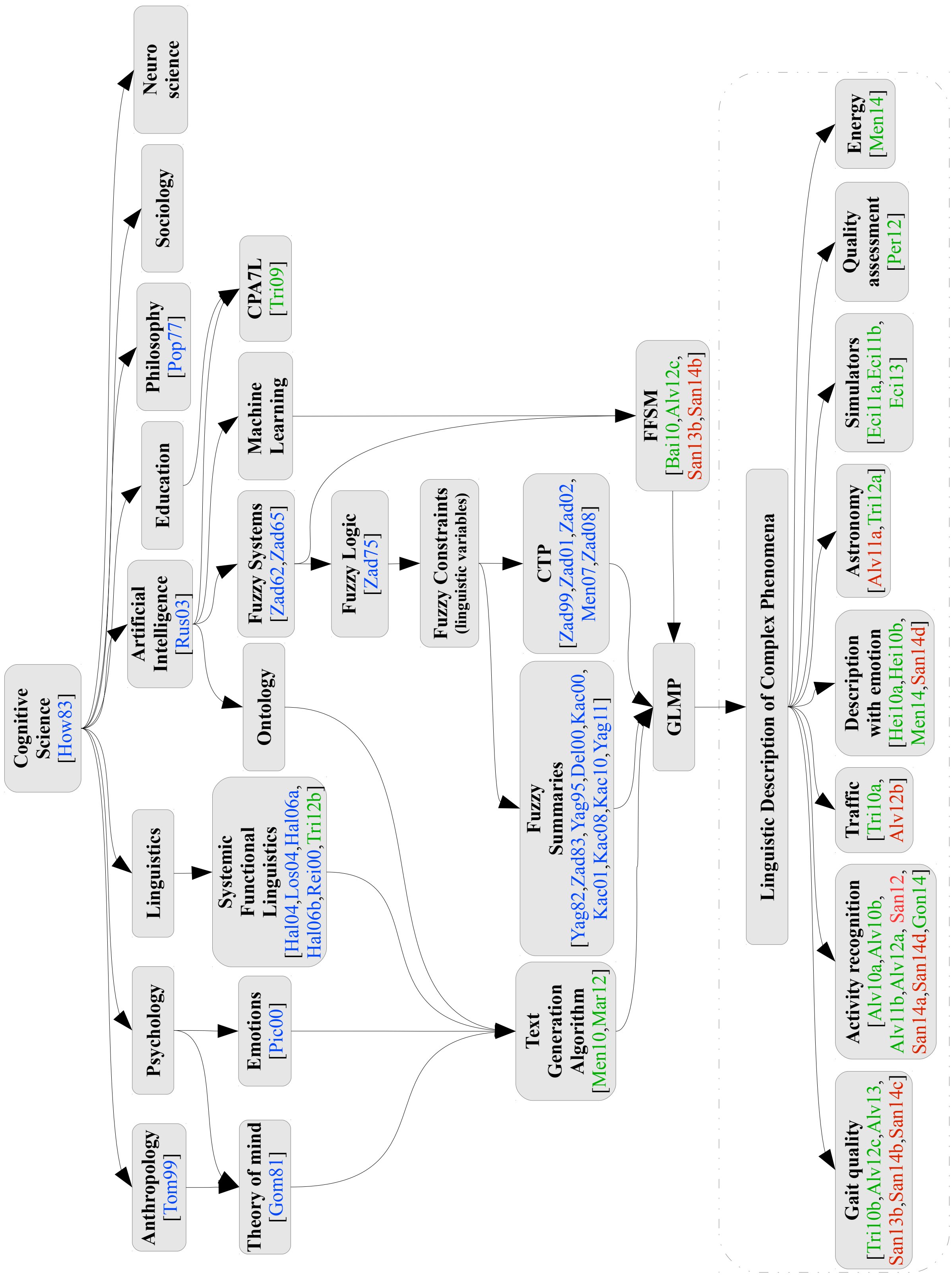


Figura 1.4: Representación gráfica de la influencia que tienen las diferentes disciplinas de la Ciencia Cognitiva en la descripción lingüística de fenómenos complejos.

Capítulo 2

Objetivos

El objetivo principal de esta tesis es el desarrollo de nuevos modelos computacionales que interpreten y describan lingüísticamente un fenómeno, de la misma manera que lo haría un ser humano. Este objetivo supone continuar con la línea de investigación iniciada en la unidad de Computación con Percepciones del ECSC, centrada en el modelado lingüístico de fenómenos complejos. El reto más importante consiste en el desarrollo de nuevos conceptos teóricos avalados por la introducción de mejoras en la resolución de problemas prácticos de diferente naturaleza.

La siguiente enumeración recoge los objetivos establecidos en el plan de investigación al comienzo de la tesis, cuyo cumplimiento fue necesario para considerar satisfactoria su finalización:

O1. **Introducir nuevos conceptos** en los distintos elementos que participan en el modelado lingüístico de fenómenos complejos (CP, PM, GLMP). Estos nuevos conceptos permiten modelar de manera más eficaz el "desconocimiento" o "no interpretabilidad" de los datos, así como recoger la información más relevante que debe aparecer en el informe final.

O2. **Ampliar la arquitectura general** del modelado lingüístico, contribuyendo de esta manera al desarrollo de la Teoría Computacional de Percepciones.

O3. **Mejorar y evolucionar el modelado de las FFSMs** con el propósito de abordar de manera más eficiente el tratamiento de fenómenos que evolucionan a lo largo del tiempo de forma quasi-periódica.

O4. **Resolver experimentos prácticos** que resuelvan problemas reales y permitan validar las contribuciones teóricas presentadas en O1, O2 y O3.

O5. **Difundir la labor investigadora** a través de la asistencia a congresos y conferencias, así como mediante la publicación de, al menos, tres artículos en revistas internacionales incluidas en el *Science Citation Index*.

Capítulo 3

Discusión de los resultados

En este capítulo se explican los resultados obtenidos durante el desarrollo de la tesis, demostrando que se han cumplido todos los objetivos expuestos en el Capítulo 2.

Puesto que esta tesis doctoral se ha desarrollado bajo la modalidad de compendio de publicaciones, este capítulo recorre brevemente cada una de las contribuciones realizadas en los artículos publicados. Para obtener una información más detallada sobre cada una de ellas, recomendamos al lector que acuda a los artículos referenciados, recogidos todos ellos en la Parte II. Publicaciones.

3.1. Introducción de nuevos conceptos en los distintos elementos que participan en el modelado de fenómenos complejos

a) La definición presentada en el apartado 1.1.1 define los elementos de una CP como un vector de expresiones lingüísticas (A) y un vector de grados de validez (W) asociados a cada una de esas expresiones lingüísticas, respectivamente. En esta tesis introducimos la multidimensionalidad del dominio de la existencia de una CP, definiendo cada uno de sus elementos como matrices multidimensionales, y entendiendo la definición existente hasta el momento como un caso particular de la misma.

La arquitectura general presentada en el apartado 1.1.5 muestra el sistema computacional propuesto para la generación automática de descripciones lingüísticas de datos. Existen gran cantidad de aplicaciones en las que el contenido de estas descripciones o informes finales varían en función de los propios datos analizados, de los objetivos perseguidos y/o del receptor del informe. De esta forma, presentamos y desarrollamos el concepto de "relevancia" como un componente más de las CPs, que se utiliza para escoger las expresiones lingüísticas más relevantes que describen el fenómeno analizado.

Tanto la multidimensionalidad de los elementos de las CPs, como el concepto de "relevancia" (R) como un elemento más de éstas, se introducen en [San 13a]. A partir de ese trabajo, una CP se presenta como una tupla de tres componentes (A, W, R), en donde A

y W son dos componentes que se han explicado en el apartado 1.1.1 y R consiste en una matriz multidimensional de grados de relevancia asignados a cada elemento de A, cuyos valores están en el intervalo $[0, 1]$. Estos valores de R representan la importancia relativa de cada expresión lingüística en el informe final y los establece el diseñador en la fase de construcción del GLMP.

Gracias a la multidimensionalidad de las CPs podemos elaborar expresiones lingüísticas más complejas que describan cada percepción. En [San 13a] se definen percepciones cuya plantilla lingüística T es, por ejemplo, *"The image contains {zero | one | two | three | four | various | many} {small | medium | big} circles"*.

b) Por otra parte, existen muchas situaciones en las que, debido a la limitada experiencia personal, el diseñador no dispone de todos los recursos necesarios para reconocer y describir los detalles más relevantes de un fenómeno. Para comunicar esa incertidumbre producida por la falta de experiencia, los seres humanos empleamos expresiones como *"por lo que yo sé, ..."*, *"parece que..."*, *"puede ser debido a..."*, etc., destacando así la falta de un conocimiento completo sobre el significado de los datos analizados. En general, cuando monitorizamos un fenómeno complejo, solo podemos reconocer los datos de entrada con un determinado grado de certeza. Como contribución en este campo, hemos enriquecido la descripción lingüística de fenómenos para superar esta limitación. Para ello, hemos creado las CPs de datos no interpretables. Entendemos por datos no interpretables aquellos que no se adaptan perfectamente al modelo disponible del fenómeno. Esta falta de conocimiento completo es una característica típica de los sistemas basados en conocimiento experto, incluyendo los sistemas basados en reglas que combinan conocimiento experto con conocimiento inducido. En aplicaciones automáticas, como son los controladores *fuzzy*, el sistema de reglas ha de ser completo y cubrir todo el ámbito de aplicación. Sin embargo, en aplicaciones que contemplan la interacción con humanos, la interpretabilidad se aprecia más que la completitud. En este sentido, cuanto mayor sea la complejidad del sistema, es decir, cuantas más reglas tenga, menor será la interpretabilidad del mismo.

El concepto de CPs de datos no interpretables se introduce en [San 14]. Para ello, extendemos su definición añadiendo la expresión lingüística a_0 en el vector A, en caso de que sea una matriz unidimensional. Esta expresión a_0 refleja aquellas situaciones en las que la percepción no se ajusta perfectamente a ninguna de las otras expresiones disponibles $(a_1, a_2, \ldots, a_n)$. Un ejemplo de expresión lingüística a_0 podría ser *"El modelo disponible no puede explicar completamente la situación actual"*. Su objetivo es modelar la falta de una interpretabilidad completa del fenómeno, proporcionando información de aquellas partes de la señal que no pueden ser explicadas por el modelo. Cuando el dominio completo que recoge una CP está bien definido, no nos preocupamos por los datos no interpretables, ya que no existen. Sin embargo, en la mayoría de las aplicaciones reales no es habitual tener un conocimiento completo del fenómeno, por lo que modelar la no interpretabilidad es esencial. Además, en muchas situaciones, este desconocimiento o "no interpretabilidad" de los datos de entrada se debe a anomalías o fallos externos al funcionamiento normal del sistema, como por ejemplo, fallos en un sensor que deja de capturar datos o su lectura es errónea. Por supuesto, la expresión lingüística a_0 tiene asociado un grado de validez w_0 dentro de la matriz W y un grado de relevancia r_0 dentro de la matriz R.

3.2. Ampliación de la arquitectura general del modelado lingüístico

a) La arquitectura inicial de un sistema computacional para la generación de descripciones lingüísticas de datos se presentó en el apartado 1.1.5. Durante el desarrollo de esta tesis doctoral se avanzó en el diseño de dicha arquitectura, redefiniendo sus etapas principales y ampliando su contenido con nuevos módulos y estructuras de datos.

En [San 13a] podemos ver una primera identificación de las dos etapas principales que encontramos en el proceso de generación de texto, llamadas "Proceso de construcción off-line" y "Proceso de instanciación on-line" (Fig. 3.1). Durante el proceso de construcción off-line, tras analizar el dominio de la aplicación y los requisitos del usuario, el diseñador recoge el conjunto de expresiones en lenguaje natural (*corpus*) que se utilizan típicamente para describir las características más relevantes del fenómeno estudiado. Asimismo, analiza el significado particular de cada expresión lingüística en cada tipo de situación y diseña tanto el GLMP como el *Report Template*. Por otra parte, durante el proceso de instanciación on-line, el sistema instancia los datos de entrada obtenidos por el módulo de adquisición de datos para generar el informe lingüístico final.

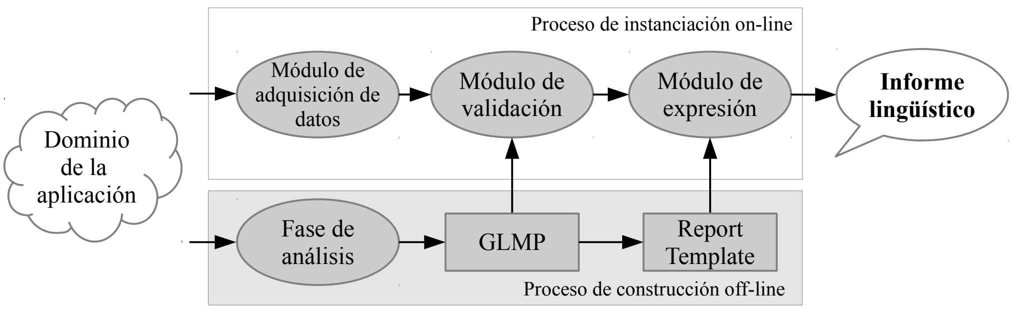


Figura 3.1: Arquitectura general propuesta en [San 13a].

b) En [San 15a] ampliamos la arquitectura añadiendo dos módulos nuevos de procesamiento: el módulo de relevancia y el módulo de expresión dinámico (Fig. 3.2). A continuación se explican en detalle las funciones de cada uno de ellos:

- El módulo de relevancia distingue automáticamente la información relevante y la irrelevante de acuerdo a las necesidades del usuario. En el apartado anterior (3.1), explicamos cómo el diseñador establece los valores de R durante el proceso off-line o construcción del GLMP. En este caso, se amplia la definición de PM, definiéndolo como una tupla (U, y, f, g, T) donde la función f es implementada por el módulo de validación (en el apartado 1.1.2 estaba representada como g) y la nueva función g, que calcula automáticamente la relevancia de cada expresión lingüística, es implementada por el módulo de relevancia. De esta manera, en función de la aplicación y de su complejidad, existirán expresiones lingüísticas cuya relevancia se pueda definir a priori durante el proceso off-line y otras cuya relevancia dependa de los resultados obtenidos. En [San 15a] se combinan ambos tipos de casos, presentando un conjunto de CPs cuyas matrices de relevancia R vienen definidas de antemano por el diseñador y otras cuyos valores dependen de los grados de validez calculados en W.

- El módulo de expresión dinámico genera los informes de manera dinámica, ajustándose en cada situación particular a los grados de validez y relevancia de las expresiones que describen el fenómeno. De acuerdo con las cuatro máximas de Grice, un informe de calidad debe contener información verdadera y relevante, expresada de manera clara y con la extensión adecuada. El objetivo perseguido con el módulo de expresión dinámico es abordar el último de esos cuatro requisitos, incorporando únicamente la información que sea cierta y relevante, ofreciendo al lector diferentes niveles de detalle y ajustando, de esta manera, la extensión del informe final.

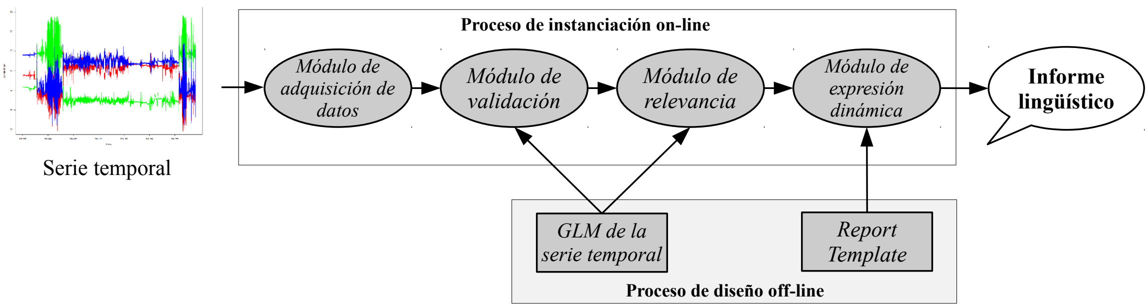


Figura 3.2: Arquitectura general propuesta en [San 15a].

c) En aquellas aplicaciones en las que se pretende conseguir un cambio de conducta o influir en el comportamiento del usuario, es importante expresar los resultados con la emoción adecuada, aportando algo más que una descripción de los resultados. En [San 15c] se amplía la arquitectura para dotar de contenido emocional el informe final. Para ello, de acuerdo con los requisitos específicos del usuario, el diseñador modela el estado emocional de un avatar, o ser virtual, por medio de un GLMP emocional (Fig. 3.3).

Este nuevo GLMP recibe como entrada la salida del GLMP clásico que, en este artículo, describe la actividad física del usuario en el instante actual, a lo largo del día o durante la semana. El avatar es el encargado de comunicar el contenido del informe final. Su estado emocional, reflejado en las expresiones lingüísticas que emplea y en sus gestos faciales, cambia dependiendo de los datos de entrada. El GLMP que modela las emociones del avatar se diseñó basándose en algunos de los modelos existentes en la literatura, como son la rueda de Plutchik, el modelo circumplejo de Russell o el modelo de Whissell. El modelo

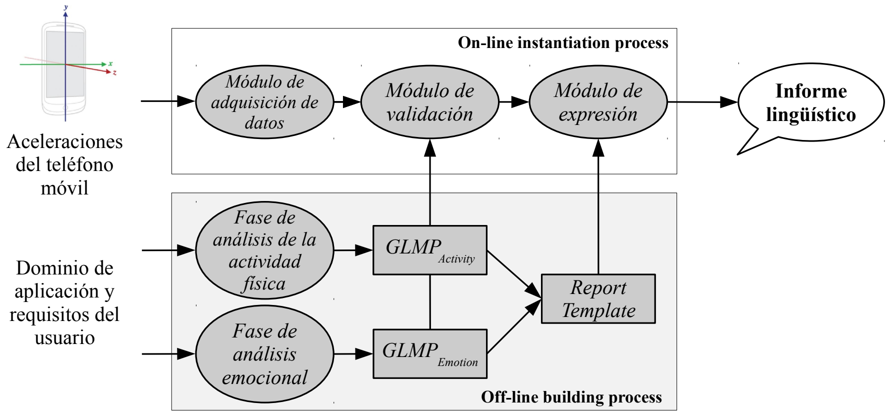


Figura 3.3: Arquitectura general propuesta en [San 15c].

diseñado selecciona una de entre nueve emociones situadas de forma circular alrededor de un punto central, que representa el estado neutral, basándose en la evaluación y la activación de los datos de entrada (Fig. 3.4). En nuestro modelo, la evaluación representa la desviación de los datos analizados con respecto al objetivo establecido. Planteamos tres posibilidades: que el objetivo se cumpla (evaluación positiva), que no se cumpla pero la desviación sea razonable (evaluación cero) o que no se cumpla ampliamente (evaluación negativa). Por su parte, la activación representa la intensidad con la que el avatar expresará los resultados, basándose en su tendencia. De este modo, con independencia de si se cumple el objetivo o no, si los resultados han mejorado la activación será pasiva (actitud del avatar más relajada), si se mantienen la activación será cero y si han empeorado la activación será activa (actitud del avatar más enérgica). Para más información sobre el modelo emocional propuesto, recomendamos al lector acudir al Capítulo 7.4.

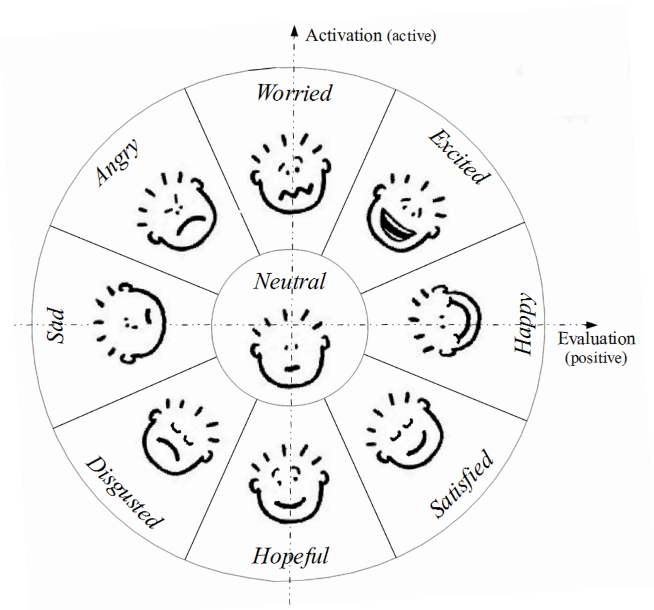


Figura 3.4: Estados emocionales propuestos en [San 15c].

3.3. Evolución en el modelado de las FFSMs

a) En lo referente a las FFSM, como un caso particular de PM de agregación, la “no interpretabilidad de los datos” presentada en el apartado 3.1 se refleja en la aparición del estado q_0 dentro del vector de estados Q, siendo el estado inicial del sistema. La Fig. 3.5 muestra un ejemplo de diagrama de estados que incorpora el estado q_0.

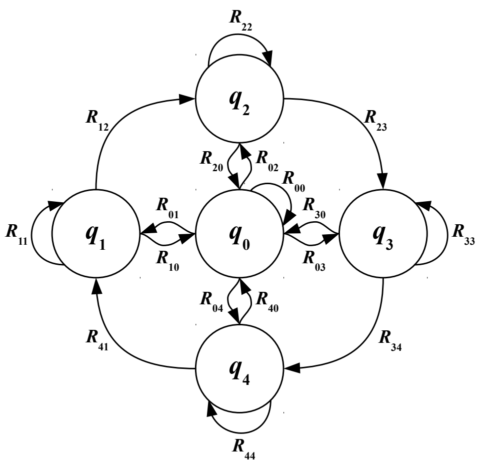


Figura 3.5: Diagrama de estados con q_0.

b) Entendemos por fenómenos quasi-periódicos aquellos cuyas señales están formadas por patrones que se repiten en el tiempo, pero que pueden presentar variaciones en el periodo y en la amplitud. Dos ejemplos de este tipo de fenómenos son los electrocardiogramas y las vibraciones producidas por los instrumentos musicales a la hora de generar sonido. En [San 13b] y [San 15b] se extiende el modelado de las señales quasi-periódicas con FFSMs, aumentando el nivel de granularidad de las descripciones lingüísticas obtenidas. El principal objetivo consiste en proporcionar un informe más detallado sobre el fenómeno quasi-periódico monitorizado, que permita entender con más claridad su comportamiento a lo largo del tiempo. Con el modelo propuesto, podemos describir lingüísticamente la evolución temporal de este tipo de señales, siendo esta descripción especialmente interesante en aquellos instantes en los que el modelo no se ajusta completamente al fenómeno. En estos casos, el sistema computacional desarrollado describe las razones que provocan esa situación, proporcionando una valiosa herramienta para monitorizar de manera más precisa la señal de entrada. Así, por ejemplo, el modelo puede devolver descripciones como *“Actualmente el modelo no reconoce la señal de entrada. La amplitud es positiva, está incrementando, y el tiempo que lleva en el estado q_1 es demasiado largo para permanecer en él, por lo que debería haber cambiado al estado q_2”*. Este análisis resulta adecuado cuando se trabaja en procesamiento de señales, control predictivo o tareas de monitorización.

El aumento del nivel de granularidad se consigue analizando en detalle el funcionamiento interno de la FFSM empleada para monitorizar la señal, proporcionando un modelo altamente interpretable, comprensible y eficiente.

Gracias a la incorporación del estado q_0 en la monitorización de fenómenos quasi-periódicos, podemos conocer el grado de similitud existente entre la señal analizada y la esperada por el modelo. De esta manera, definimos el concepto de *matching degree* (μ_U), que se calcula como:

$$\mu_U = 1 - \frac{\sum_{t=1}^{n} w_{q_0}[t]}{n} \tag{3.1}$$

donde w_{q_0} representa los grados de validez del estado q_0 a lo largo de un periodo de n muestras. El coeficiente μ_U toma valores en el rango [0,1], siendo igual a uno si el estado "no interpretable" no se ha activado, cero cuando ha estado activado durante todo el tiempo y un número real entre cero y uno para el resto de casos. Por ejemplo, si monitorizamos el comportamiento de la señal senoidal $u(t) = sin(10t)$ a lo largo del tiempo, podemos identificar los cuatro estados q_1, q_2, q_3 y q_4 representados en la Fig. 3.6.

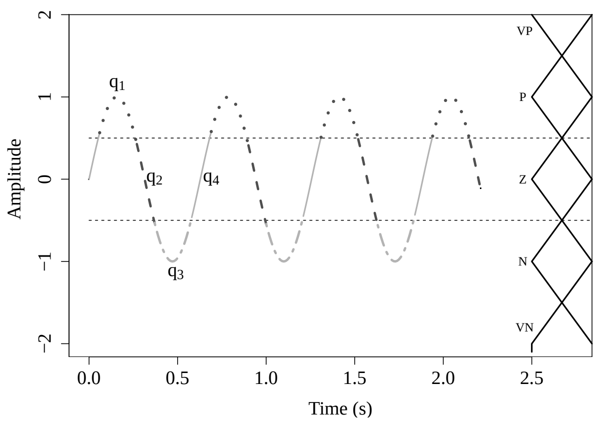


Figura 3.6: Señal senoidal con cinco etiquetas lingüísticas en amplitud, que definen los estados q_1, q_2, q_3 y q_4.

Las Figs. 3.7, 3.8, 3.9 y 3.10 muestran la evolución de estos estados, además del estado q_0, conforme la señal se aleja o se asemeja al modelo previsto. Cuanto más interpretable sea la señal, mayor será el *matching degree*, y viceversa.

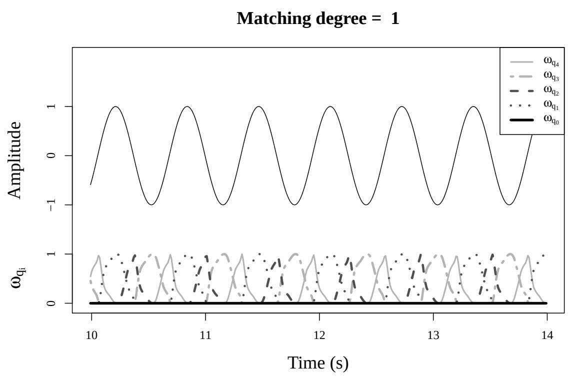


Figura 3.7: Estados y μ_U cuando la señal es exactamente la esperada.

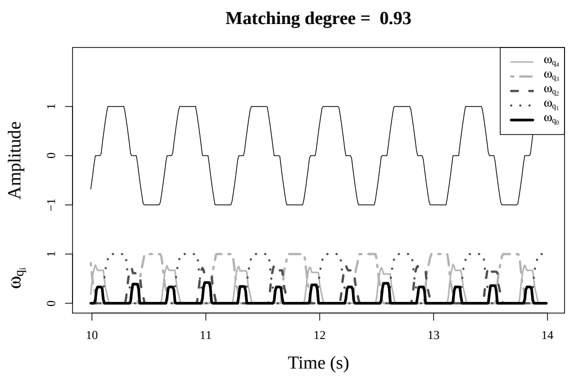


Figura 3.8: Estados y μ_U cuando la señal es ligeramente diferente de la esperada.

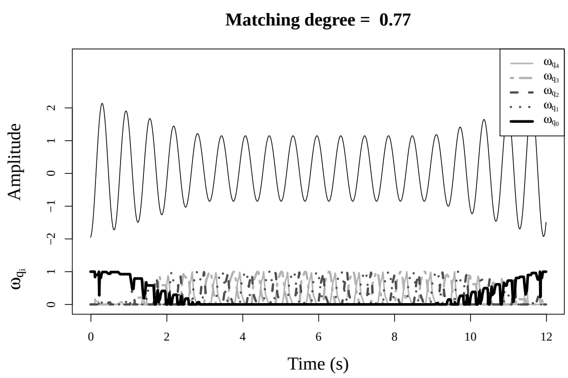


Figura 3.9: Estados y μ_U cuando la señal varía su amplitud a lo largo del tiempo.

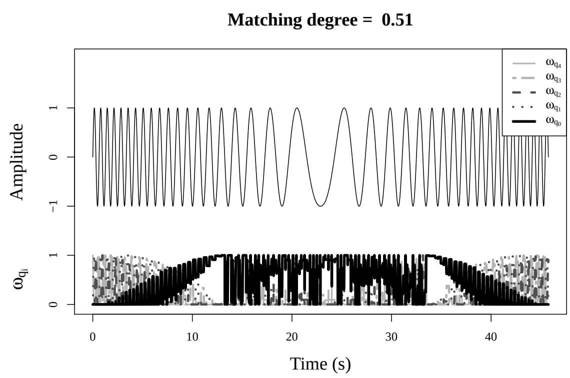


Figura 3.10: Estados y μ_U cuando la señal varía su periodo a lo largo del tiempo.

3.4. Resolución de casos prácticos que validen las contribuciones teóricas

Cada una de las contribuciones teóricas presentadas en 3.1, 3.2 y 3.3 han sido validadas a partir de la resolución práctica de problemas reales. Estos casos prácticos se han dividido en los tres campos de aplicación que se describen a continuación:

3.4.1. Análisis de las estructuras circulares presentes en la superficie de Marte

Fruto de la colaboración con el Instituto Nacional de Tecnología Aeroespacial (INTA), en [San 13a] se desarrolló un sistema computacional que describe lingüísticamente las estructuras circulares presentes en la superficie de Marte (Fig. 3.11).

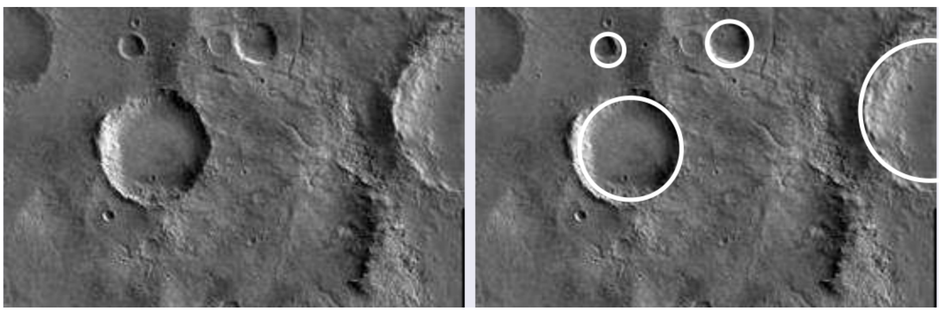


Figura 3.11: Ejemplo de la detección de círculos en una imagen.

Los satélites que orbitan sobre Marte capturan diariamente un gran volúmen de imágenes de su superficie, y proporcionan a los expertos información geológica importante. Entre otros, el hallazgo y análisis de estructuras con forma circular resulta interesante puesto que su origen se puede relacionar con la presencia de actividad volcánica, aunque también con simples impactos de meteoritos. El principal problema reside en la incapacidad humana para analizar las miles de imágenes que se capturan cada día, por lo que es necesario diseñar un sistema que automáticamente analice y describa la información más relevante encontrada, omitiendo la irrelevante. Los informes generados son estudiados en detalle por los expertos, que ven como se reduce de manera importante el estudio total de casos a los más relevantes.

Algunos ejemplos de las descripciones obtenidas para dos imágenes diferentes son:

"La imagen contiene dos círculos grandes y dos medianos. Los dos círculos grandes están ubicados en la parte central inferior y superior de la imagen"

"La imagen no contiene ningún círculo grande pero sí uno mediano"

Tal y como puede observarse en los ejemplos anteriores, el contenido y extensión de las descripciones varían en función del resultado obtenido en el análisis de la imagen, según la teoría presentada en el apartado 3.2. Así, en este caso práctico, cuando se detecta la presencia de algún círculo grande es relevante incluir en el informe su posición. Esta selección de información la realiza el módulo de expresión a partir de un *Fuzzy tree of choices* y de la combinación de los grados de validez y relevancia de cada sentencia (para una información más detallada recomendamos acudir al Capítulo 5.1).

3.4.2. Análisis de la actividad humana

Una de las principales líneas de investigación de esta tesis doctoral se basa en el desarrollo de un sistema computacional que analiza las actividades físicas más comunes de una persona. La monitorización de la actividad física es especialmente útil en el ámbito de la asistencia médica personalizada, tanto para la identificación de potenciales patrones patológicos en personas con tendencia a padecer trastornos mentales (depresión o trastorno bipolar), como en la detección del grado de inactividad, siendo ésta una de las causas que producen o agravan enfermedades cardiovasculares, óseas, crónicas, etc. Su estudio resulta de gran ayuda para profesionales (endocrinos, psicólogos, cardiólogos, ...) que pretenden ajustar tratamientos y analizar la evolución física de sus pacientes.

Nuestra primera contribución en este campo de investigación se centra en el proceso de captura de los datos. Hasta hace poco, la manera que tenía un médico para evaluar la actividad física de un paciente era realizarle cuestionarios que estaban condicionados por su memoria y también por su subjetividad a la hora de sincerarse. Para resolver este problema y proporcionar un método de captura de información más objetivo, durante los últimos años han ido apareciendo gran cantidad de sistemas de reconocimiento de la actividad que recogen y analizan las aceleraciones producidas durante el movimiento. Algunas de estas soluciones requieren la colocación de diferentes sensores distribuidos por el cuerpo del usuario, resultando incómodo e imprácticable a la hora de realizar una monitorización continua. Otras soluciones más sofisticadas se centran en colocar un único sensor, y gracias al potencial de las nuevas tecnologías, durante los últimos años se están aprovechando los acelerómetros incorporados en los actuales teléfonos móviles para capturar las aceleraciones producidas por el movimiento. En [San 12] se propone un modelo que describe lingüísticamente las principales actividades de una persona (estar parado, sentado, caminando, corriendo, subiendo y bajando escaleras...) mediante el análisis de las aceleraciones producidas en su teléfono móvil mientras se lleva en el bolsillo del pantalón. La Fig. 3.12 muestra las aceleraciones recogidas por el teléfono móvil durante un periodo de grabación diaria y la Fig. 3.13 muestra la secuencia de actividades identificadas tras analizar su evolución.

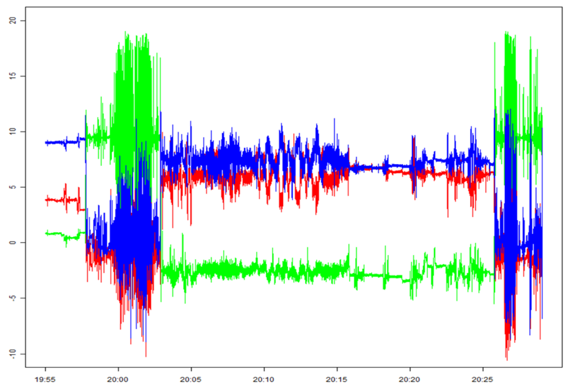


Figura 3.12: Aceleraciones en los ejes x, y, z durante un periodo de tiempo.

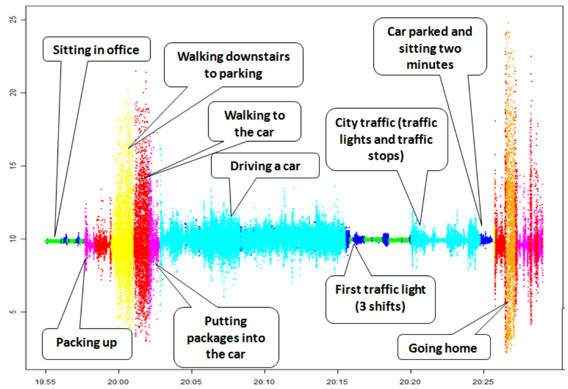


Figura 3.13: Identificación de las actividades físicas de una persona.

Con el objetivo de facilitar lo máximo posible la captura de datos, en [San 15a] se da un paso más, permitiendo que el teléfono móvil se lleve donde el propio usuario estime oportuno, incluso en un bolso de mano o mochila. Con este avance se reduce el número de actividades reconocidas pero, a cambio, conseguimos eliminar la gran limitación de los modelos existentes de exigir una ubicación fija del dispositivo. El modelo propuesto es capaz de reconocer cuando el usuario no lleva el teléfono móvil encima, teniendo en cuenta esta información a la hora de estimar el nivel de actividad física diario. Para ello, el módulo de adquisición de datos recoge como entrada las aceleraciones a_x, a_y y a_z producidas en el teléfono móvil (Fig. 3.12). Estas aceleraciones se procesan para obtener su módulo (ρ) y su desviación típica (σ) en ventanas temporales de 20 segundos. Es importante destacar que ρ y σ son dos medidas invariantes con respecto de la orientación y posición del teléfono móvil. En este sentido, la evolución de a_x, a_y y a_z por separado no resulta interesante, pero sí la magnitud y variabilidad de su combinación, lo cual está representado por ρ y σ. La forma de calcularlas es la siguiente:

$$\rho = \sqrt{a_x^2 + a_y^2 + a_z^2} \tag{3.2}$$

$$\sigma = \sqrt{\frac{\sum_{i=1}^{N}(\rho_i - \bar{\rho})^2}{N-1}} \tag{3.3}$$

Otra de las aportaciones importantes de este trabajo hace relación a las contribuciones teóricas presentadas en los apartados 3.1 y 3.2. Por una parte, el modelo de la actividad física propuesto tiene en cuenta aquellos instantes en los que no se reconoce la señal de entrada, tratándolo como información "no interpretable". Este dato también será importante a la hora de estimar el nivel de actividad física diario. La Fig. 3.14 muestra el diagrama de estados de la FFSM que reconoce las actividades del usuario.

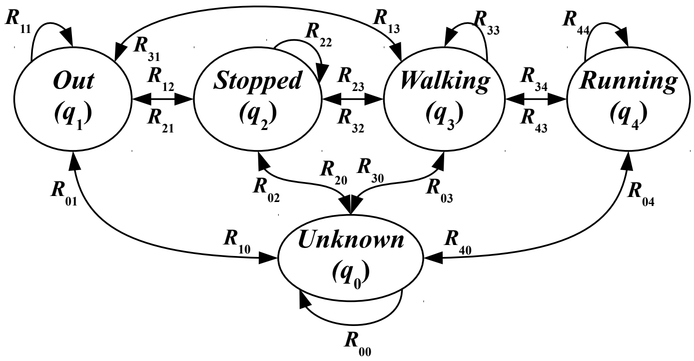


Figura 3.14: Diagrama de estados propuesto en [San 15a].

Por otra parte, con el objetivo de generar informes de calidad que cumplan las cuatro máximas de Grice, la arquitectura de este trabajo incluye el módulo de relevancia y el módulo de expresión dinámico. De esta manera, el modelo es capaz de generar informes con tres niveles de granularidad: instantánea, diaria y semanal. El contenido de estos informes se ajustará a los resultados obtenidos y su relevancia en cada contexto específico.

Por último, en [San 15c] se amplía el modelo de la actividad física, proporcionando al usuario información acerca de los requisitos energéticos diarios que necesita en función de su nivel de actividad física y su índice de masa corporal. La Fig. 3.15 muestra el GLMP diseñado para modelar la actividad física de una persona y sus requisitos energéticos. Si nos fijamos en el ejemplo de GLMP presentado en el Capítulo 1.1.3 podemos ver un cambio en la notación utilizada para designar las CPs y las PMs. La distinción entre percepciones de primer y segundo orden se pierde, y los subíndices empleados corresponden a alguna característica de la propia percepción (ρ, σ, etc.), en lugar de utilizar índices numéricos. Para una información más detallada acudir al Capítulo 7.4.

Con el objetivo de promover el *self-tracking* o auto-seguimiento de la actividad física, dotamos el informe final de contenido emocional. A la hora de motivar al usuario para que alcance el nivel de actividad física establecido por un experto o por él mismo, es imprescindible interactuar y comunicar la información relevante de manera efectiva, estableciendo un vínculo afectivo con él. Por esta razón, decidimos incluir técnicas de *Affective Computing* en nuestra arquitectura general (Fig. 3.3), añadiendo un GLMP que modela las emociones de un ser virtual. Este ser virtual será el encargado de comunicar los resultados de la manera más adecuada en cada situación. Por ejemplo, si los resultados son negativos, el avatar se mostrará preocupado, triste o incluso enfadado; si estos persisten en el tiempo

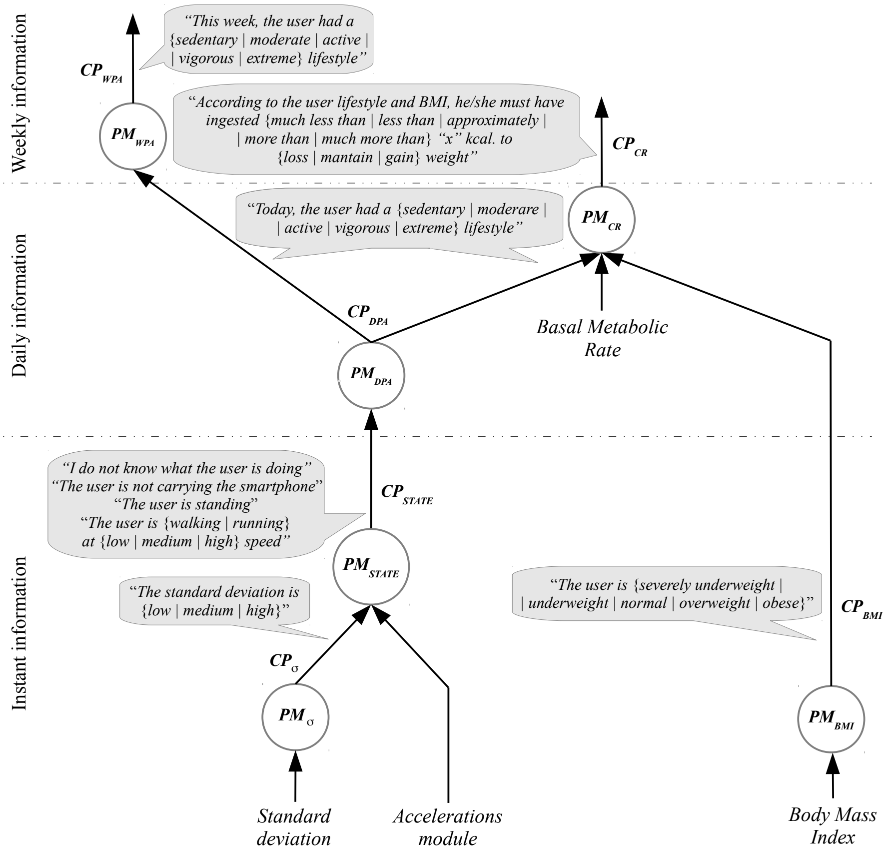


Figura 3.15: GLMP propuesto en [San 15c] para modelar la actividad física.

se mostrará disgustado; si por el contrario son positivos, se mostrará esperanzado, feliz o excitado, y si se alcanza el objetivo perseguido, se mostrará satisfecho. Tanto su expresión facial, como las expresiones lingüísticas utilizadas estarán condicionadas por su estado de ánimo.

3.4.3. Análisis de la marcha humana

Otra de las principales líneas de investigación de esta tesis ha sido la mejora y evolución del modelo de análisis de la calidad de la marcha creado y desarrollado en [Tri 10b], [Alv 13] y [Alv 12c].

Modelar la marcha humana consiste en estudiar su biomecánica con el objetivo de cuantificar los factores que gobiernan la funcionalidad de las extremidades inferiores. Un ciclo de marcha es un fenómeno quasi-periódico que se define como el intervalo comprendido entre dos eventos sucesivos, generalmente correspondientes al contacto del talón con el suelo del mismo pie. Tal y como puede apreciarse en la Fig. 3.16, la marcha humana se puede dividir en cuatro fases o estados, que son el doble soporte y el balanceo de la pierna derecha, y el doble soporte y balanceo de la pierna izquierda.

La marcha es una tarea compleja que requiere de la coordinación entre el sistema musculoesquelético y el neuronal, asegurando una dinámica correcta. Por tanto, su análi-

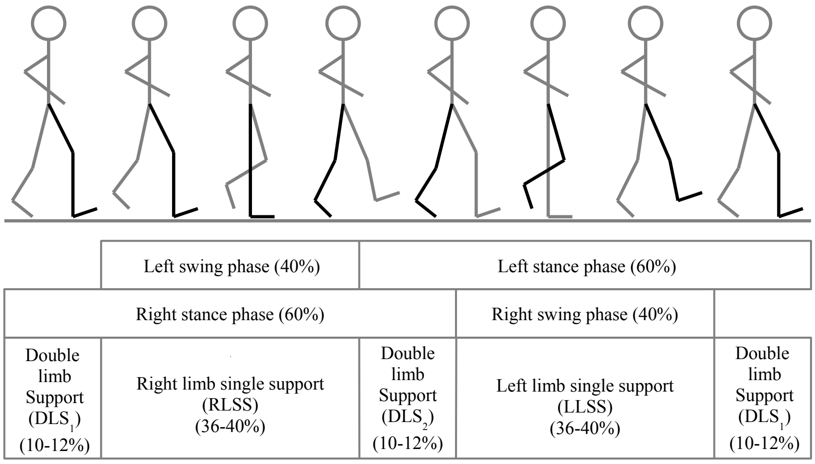


Figura 3.16: Estados y eventos existentes en un ciclo de marcha humana.

sis puede ayudar en el diagnóstico y tratamiento de desórdenes que le afectan, así como proporcionar una herramienta para la evaluación de intervenciones clínicas y programas de rehabilitación. Actualmente, realizar una valoración de la marcha humana solo es posible acudiendo al especialista para una valoración visual o acudiendo a los laboratorios de biomecánica, preparados con complejos sistemas de grabación y medición de la pisada. Con el objetivo de desarrollar una alternativa económica, precisa y cómoda para el usuario, desde la unidad de Computación con Percepciones del ECSC hemos desarrollado un modelo computacional capaz de evaluar la calidad de la marcha mediante el análisis de las aceleraciones producidas al caminar, llevando el teléfono móvil centrado en la cintura (Fig. 3.17).

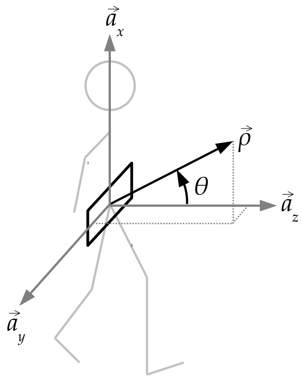


Figura 3.17: Representación gráfica de la colocación del teléfono móvil.

La primera mejora del modelo se presenta en [San 14], donde se capturan las aceleraciones producidas en el teléfono móvil del usuario, colocado centrado en la cintura. En este trabajo se introduce el análisis de la "no interpretabilidad" de los datos, presentada en el apartado 3.1, contribuyendo notablemente en el diseño de un modelo de la marcha más robusto y eficiente. La incorporación del estado q_0 dentro de la FFSM que modela la marcha humana permite identificar el mejor intervalo temporal para realizar el análisis de la señal de entrada, es decir, analizar aquellas partes de la señal a las que mejor se ajusta el modelo. De esta manera, se eliminan aquellos instantes en los que el usuario no está caminando, se ha producido algún evento extraño, etc. (Fig. 3.18). Una vez que el análisis se ha realizado, podemos proporcionar un indicador (*matching degree*) sobre el mayor o menor grado con el que el modelo ha reconocido la señal y, por tanto, estimar la confianza de los resultados generados.

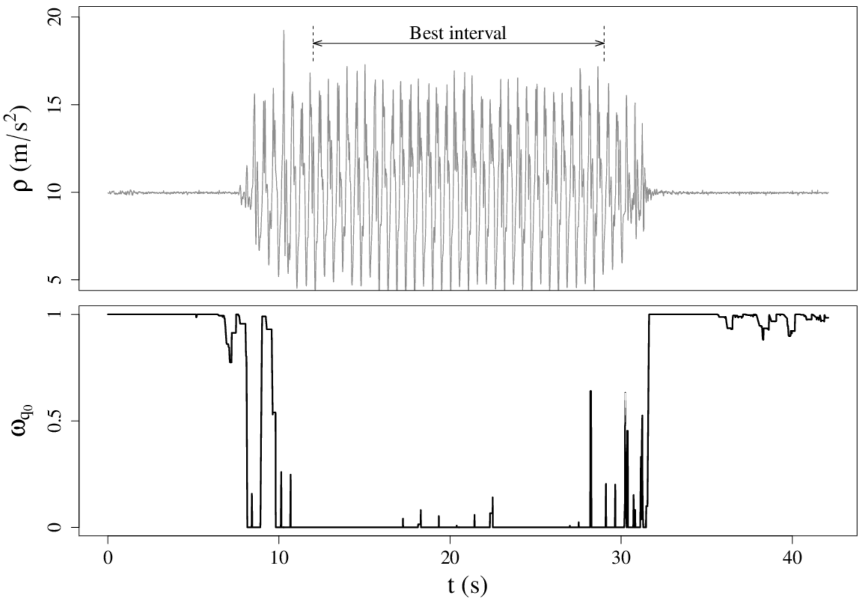


Figura 3.18: Aceleraciones producidas al caminar y selección del mejor intervalo para su análisis.

Otra mejora importante, también desarrollada en [San 14], se refiere al ajuste automático de los parámetros de amplitud y de periodo que influyen en el modelado de la marcha. Debido a que la marcha humana es un fenómeno quasi-periódico que no produce una señal totalmente constante y repetible, estos parámetros pueden variar. Las principales causas son debidas a cambios en la velocidad y en la fuerza aplicada al caminar, pero también pueden surgir importantes diferencias derivadas del tipo de calzado, la superficie sobre la que caminamos, etc. Con el objetivo de desarrollar un modelo robusto que reconozca las fases de la marcha independientemente de todas esas variaciones, en este trabajo se desarrolla un sistema de ajuste automático de los parámetros de amplitud y de periodo, modificándolos ligeramente para acomodarse a cada situación. Las Figs. 3.19 y 3.20 muestran la identificación de las fases de la marcha para una persona con marcha normal y otra con cojera en la pierna derecha, lo cual le hace dar golpes más fuertes con esta pierna y caminar más despacio (fijarse en el eje temporal). Se puede apreciar como el reconocimiento de las fases es correcto a pesar de las notables diferencias en la señal.

En [San 15b] se amplía el modelo existente, incrementando el nivel de detalle en la descripción de la marcha, de acuerdo a la contribución teórica presentada en el apartado 3.3. Además, se profundiza en el análisis de los parámetros que afectan a la marcha

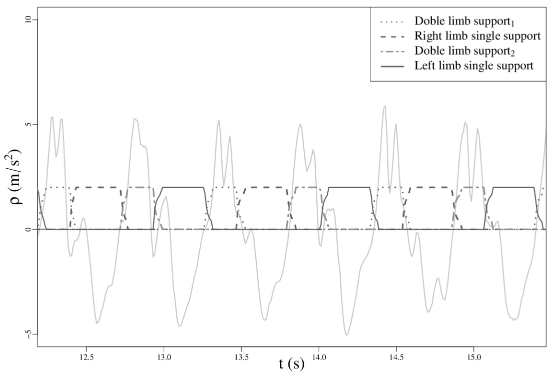


Figura 3.19: Identificación de estados para una marcha normal.

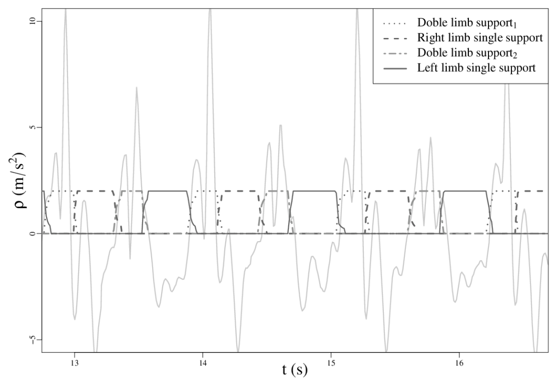
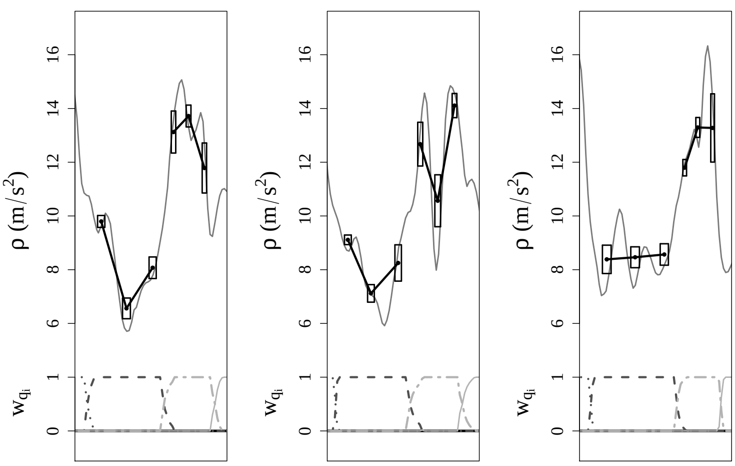


Figura 3.20: Identificación de estados para una marcha asimétrica.

humana, desarrollando un clasificador difuso basado en reglas que identifica cinco tipos de marcha: normal, de puntillas, arrastrando ambos pies, arrastrando el pie derecho o arrastando el pie izquierdo. La Fig. 3.21 muestra un ejemplo de los distintos patrones de pisada para las fases de una pierna (balanceo y doble soporte) según se camine normal, de puntillas o arrastrando.

Identificar el tipo de marcha es importante para ayudar en el diagnóstico de una enfermendad, comprobar su evolución y analizar el efecto de una operación, un proceso de rehabilitación o una medicación. En particular, la importancia de detectar que una persona arrastra un pie o camina de puntillas se ha destacado en multitud de estudios debido a que un contacto deficiente con el suelo puede provocar una caída. El modelo desarrollado proporciona un 84 % de precisión en la clasificación, encontrándose a la altura de algoritmos de clasificación clásicos, como las redes neuronales o los *Support Vector Machine*, pero con la ventaja de que permite una alta interpretabilidad de los resultados obtenidos.



(a) Normal (b) De puntillas (c) Arrastrando

Figura 3.21: Patrones característicos dependiendo del tipo de marcha.

3.5. Difusión de los resultados

Cada una de las contribuciones citadas en los apartados anteriores se ha presentado a la comunidad científica a través de la publicación de artículos en revistas internacionales incluidas en el *Science Citation Index* y mediante la asistencia a congresos científicos relacionados con cada área de investigación. En particular, se han publicado dos artículos en la revista *Fuzzy Sets and Systems* [San 14, San 15a], otros dos en *Applied Soft Computing* [San 13a, San 15b] y uno en *Expert Systems with Applications* [San 15c]. Además, se ha asistido a la cuarta edición del *International Workshop of Ambient Assisted Living* (IWAAL'12) celebrado en Vitoria en 2012 [San 12], y al décimo congreso internacional *Flexible Query Answering Systems* (FQAS'13) celebrado en Granada en 2013 [San 13b].

Fruto de la investigación llevada a cabo en el ámbito de la marcha humana, durante los últimos años hemos tenido la oportunidad de colaborar estrechamente con distintos expertos en rehabilitación física. En particular, estamos colaborando con D. Ricardo Llavona Fernández, médico rehabilitador del Servicio de Salud del Principado de Asturias, especialista en medicina física y rehabilitación. Con su ayuda, hemos iniciado un proceso de experimentación en el Hospital Valle del Nalón (Asturias), monitorizando la evolución de los pacientes con artrosis de rodilla (con o sin recambio articular). A raíz de esta colaboración nos han otorgado el premio al "Mejor proyecto de investigación del área Sanitaria VIII" en Diciembre de 2014.

También estamos colaborando con D. Ignacio García Perea, fisioterapeuta en el Centro de Terapia Neurológica de Majadahonda, especialista en rehabilitación de pacientes con daños neurológicos. Gracias a él, hemos tenido la posibilidad de monitorizar la evolución de pacientes con ictus, traumatismos craneoencefálicos, Parkinson, etc., y comprobar el efecto que tiene cada una de las intervenciones del especialista en su mejora continua.

Los resultados obtenidos en esta línea de investigación son tan positivos y esperanzadores que actualmente nos encontramos en proceso de creación de la primera spin-off del ECSC, centrada en diseñar y desarrollar aplicaciones que evalúen la calidad de la marcha de personas con distintos tipos de patologías. Para llevar a cabo esta acción ha sido necesario formarse adecuadamente en materia relacionada con el emprendimiento. En especial, he recibido un curso acelerador de emprendimiento en el "Programa de impulso a la creación de empresas innovadoras en el campo de la Salud (especialización eHealth)" organizado por el CEEI (Centro Europeo de Empresas e Innovación) del Principado de Asturias.

Finalmente, en paralelo a la realización de esta tesis doctoral, he complementado las tareas de investigación con la participación en proyectos que ponen en valor la tecnología desarrollada. En particular, en colaboración con la empresa EDP Energía se ha desarrollado el proyecto ELENNA (Electricidad en LENguaje NAtural), que ha consistido en el diseño de un sistema computacional que genera descripciones lingüísticas personalizadas sobre el consumo eléctrico de cada cliente. La comunicación con cada usuario corre a cargo de un agente virtual cuyo estado emocional depende de los resultados obtenidos.

En esta misma dirección, desde el pasado mes de enero, el ECSC participa en el proyecto europeo NATCONSUMERS (NATural language energy for promoting CONSUMER Sustainable behaviour), que tiene como objetivo ayudar a reducir un 20 % el consumo eléctrico doméstico de la Unión Europea.

Capítulo 4

Conclusiones y trabajo futuro

Durante esta tesis doctoral se ha continuado la línea de investigación iniciada en la unidad de Computación con Percepciones del *European Centre for Soft Computing* en el modelado lingüístico de fenómenos complejos. Para ello, se han desarrollado nuevos conceptos que mejoran la interpretabilidad, robustez y aplicabilidad de los modelos existentes. Las principales contribuciones teóricas han ido dirigidas a la introducción de nuevos elementos, así como a la ampliación de la arquitectura general a través de nuevos módulos y estructuras de datos. En particular, se ha incorporado el módulo de relevancia, que calcula automáticamente la importancia de tiene cada expresión lingüística a la hora de incluirla en la descripción del fenómeno analizado, ajustando el contenido a cada situación en particular. Además, el módulo de expresión dinámico permite adaptar el nivel de detalle del informe final según el receptor del mensaje, aumentando la calidad de las descripciones generadas según las máximas de Grice. En aquellas aplicaciones en las que se pretende conseguir un cambio de conducta o influir en el comportamiento del usuario, es muy importante expresar los resultados con la emoción adecuada. Así, se ha desarrollado un modelo emocional que ajusta el tono y la intención con los que se expresa la información, estableciendo un vínculo emocional con el usuario y aumentando la efectividad de los mensajes.

Por otra parte, con el objetivo de proporcionar un informe lo más detallado posible sobre el fenómeno estudiado, se han creado las CPs de datos no interpretables, que modelan aquellas situaciones en las que no se posee un conocimiento completo de la señal. Esta mejora nos permite extraer información de aquellos instantes en las que el modelo no encaja perfectamente con la señal, proporcionando una herramienta valiosa en la monitorización de la misma. Todos estos avances teóricos han contribuido en la evolución de las *Fuzzy Finite State Machines*, consiguiendo, en particular, modelar más eficientemente los fenómenos quasi-periódicos.

Con el fin de avalar todas y cada una de las contribuciones teóricas descritas anteriormente, se ha abordado el estudio y la resolución de problemas prácticos con demanda en el mundo real, comprobando de esta manera la efectividad y potencialidad de las mejoras y avances introducidos. Estos desarrollos prácticos se han llevado a cabo en colaboración con empresas e instituciones que han puesto en valor nuestra tecnología. En especial, hemos mejorado ampliamente el modelo de la marcha humana, despertando gran interés en el sector sanitario, cuyo reconocimiento se ha traducido en forma de premios y colaboraciones. Además, los prometedores resultados cosechados nos han impulsado a contemplar la creación de la primera spin-off del *European Centre for Soft Computing*.

A pesar de la mejora cualitativa en el proceso de generación de informes, creemos que existen aún muchas oportunidades de mejora en el mundo del modelado lingüístico. En particular, la recopilación del *corpus* de expresiones comúnmente utilizadas para describir el fenómeno analizado, así como las técnicas de generación del informe final, admiten un largo recorrido de investigación en trabajos futuros.

A corto plazo, se pretenden orientar las herramientas existentes para el tratamiento de nuevos problemas específicos. Por ejemplo, en el campo del análisis de la actividad queremos profundizar en el desarrollo de una herramienta de apoyo para personas con tendencia a padecer desórdenes mentales, como la depresión y el trastorno bipolar. De igual manera, el modelo de la marcha humana se especializará en el análisis de patologías concretas que afectan a la marcha, como son algunas enfermedades neurológicas y musculoesqueléticas.

Finalmente, podemos concluir que los resultados obtenidos demuestran que se ha cumplido el objetivo principal de contribuir de forma teórica y práctica en el modelado lingüístico de fenómenos complejos. Además, se han alcanzado todos y cada uno de los distintos objetivos definidos al comienzo de la misma, resumidos anteriormente. El éxito del trabajo realizado se refleja tanto en las publicaciones en revistas cientícas como en los proyectos desarrollados durante los últimos cuatro años.

Bibliografía

[Alo 07] J. M. Alonso. *Interpretable fuzzy systems modeling with cooperation between expert and induced knowledge*. PhD thesis, Technical University of Madrid, http://oa.upm.es/588/, 2007.

[Alo 08] J. M. Alonso, L. Magdalena, and S. Guillaume. "HILK: A new methodology for designing Highly Interpretable Linguistic Knowledge bases using the fuzzy logic formalism". *International Journal of Intelligent Systems*, Vol. 23, No. 7, pp. 761–794, July 2008.

[Alo 11a] J. M. Alonso and L. Magdalena. "HILK++: an interpretability-guided fuzzy modeling methodology for learning readable and comprehensible fuzzy rule-based classifiers". *Soft Computing*, Vol. 15, No. 10, pp. 1959–1980, 2011.

[Alo 11b] J. M. Alonso and L. Magdalena. "Special issue on interpretable fuzzy systems". *Information Sciences*, Vol. 181, No. 20, pp. 4331–4339, 2011.

[Alv 10a] A. Alvarez-Alvarez, J. M. Alonso, G. Trivino, N. Hernández, F. Herranz, A. Llamazares, and M. Ocaña. "Human Activity Recognition applying Computational Intelligence techniques for fusing information related to WiFi positioning and body posture". In: *Proceedings of the 2010 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), Barcelona, Spain*, pp. 1881–1885, July 2010.

[Alv 10b] A. Alvarez-Alvarez, J. M. Alonso, G. Trivino, N. Hernández, F. Herranz, and M. Ocaña. "Towards people indoor localization combining WiFi and human motion recognition". In: *Actas XV Congreso Español sobre Tecnologías y Lógica Fuzzy (ESTYLF), Huelva, Spain*, pp. 7–12, February 2010.

[Alv 11a] A. Alvarez-Alvarez, D. Sanchez-Valdes, and G. Trivino. "Automatic Linguistic Description about Relevant Features of the Mars' Surface". In: *Proceedings of the 11th International Conference on Intelligent Systems Design and Applications (ISDA), Córdoba, Spain*, pp. 154–159, November 2011.

[Alv 11b] A. Alvarez-Alvarez, G. Trivino, and O. Cordón. "Body Posture Recognition By Means Of A Genetic Fuzzy Finite State Machine". In: *Proceedings of the 5th IEEE International Workshop on Genetic and Evolutionary Fuzzy Systems (GEFS), Paris, France*, pp. 60–65, April 2011.

[Alv 12a] A. Alvarez-Alvarez, J. M. Alonso, and G. Trivino. “Human activity recognition in indoor environments by means of fusing information extracted from intensity of WiFi signal and accelerations”. *Information Sciences*, Vol. 233, pp. 162–182, 2012.

[Alv 12b] A. Alvarez-Alvarez, D. Sanchez-Valdes, G. Trivino, Á. Sánchez, and P. D. Suárez. “Automatic linguistic report of traffic evolution in roads”. *Expert Systems with Applications*, Vol. 39, No. 12, pp. 11293–11302, 2012.

[Alv 12c] A. Alvarez-Alvarez, G. Trivino, and O. Cordón. “Human gait modeling using a genetic fuzzy finite state machine”. *Fuzzy Systems, IEEE Transactions on*, Vol. 20, No. 2, pp. 205–223, April 2012.

[Alv 13] A. Alvarez-Alvarez and G. Trivino. “Linguistic description of the human gait quality”. *Engineering Applications of Artificial Intelligence*, Vol. 26, pp. 13–23, 2013.

[Arg 13] L. Arguelles and G. Trivino. “I-struve: Automatic linguistic descriptions of visual double stars”. *Engineering Applications of Artificial Intelligence*, Vol. 26, No. 9, pp. 2083–2092, 2013.

[Bai 10] G. Bailador and G. Trivino. “Pattern recognition using temporal fuzzy automata”. *Fuzzy Sets and Systems*, Vol. 161, No. 1, pp. 37–55, 2010.

[Buc 93] J. J. Buckley. “Sugeno type controllers are universal controllers”. *Fuzzy Sets and Systems*, Vol. 53, No. 3, pp. 299–303, February 1993.

[Cao 07] Y. Cao, M. Ying, and G. Chen. “Retraction and Generalized Extension of Computing With Words”. *Fuzzy Systems, IEEE Transactions on*, Vol. 15, No. 6, pp. 1238–1250, 2007.

[Cas 03a] J. Casillas, O. Cordón, F. Herrera, and L. Magdalena, Eds. *Accuracy improvements in linguistic fuzzy modeling.* Vol. 129 of *Studies in Fuzziness and Soft Computing*, Springer, 2003.

[Cas 03b] J. Casillas, O. Cordón, F. Herrera, and L. Magdalena, Eds. *Interpretability Issues in Fuzzy Modeling.* Vol. 128 of *Studies in Fuzziness and Soft Computing*, Springer, 2003.

[Cas 95] J. L. Castro. “Fuzzy logic controllers are universal approximators”. *Systems, Man and Cybernetics, IEEE Transactions on*, Vol. 25, No. 4, pp. 629–635, April 1995.

[Del 00] M. Delgado, D. Sánchez, and M. Vila. “Fuzzy cardinality based evaluation of quantified sentences”. *Int. Journal of Approximate Reasoning*, Vol. 23, pp. 23–66, 2000.

[Eci 11a] L. Eciolaza, G. Trivino, B. Delgado, J. Rojas, and M. Sevillano. "Fuzzy linguistic reporting in driving simulators". In: *Proceedings of the 2011 IEEE Symposium on Computational Intelligence in Vehicles and Transportation Systems, Paris, France*, pp. 30–37, 2011.

[Eci 11b] L. Eciolaza and G. Trivino. "Linguistic Reporting of Driver Behavior: Summary and Event Description". In: *Proceedings of the 11th International Conference on Intelligent Systems Design and Applications (ISDA), Córdoba, Spain*, pp. 148–153, 2011.

[Eci 13] L. Eciolaza, M. Pereira-Fariña, and G. Trivino. "Automatic linguistic reporting in driving simulation environments". *Applied Soft Computing*, Vol. 13, No. 9, pp. 3956–3967, 2013.

[Gom 91] A. R. Gómez. *Objetos con mente.* Alianza Editorial, 1991.

[Gon 14] L. González-Villanueva, A. Alvarez-Alvarez, L. Ascari, and G. Trivino. "A tool for linguistic assessment of rehabilitation exercises". *Applied Soft Computing*, Vol. 14, pp. 120–131, 2014.

[Hal 04] M. A. Halliday and C. M. Matthiessen. "An introduction to functional grammar". 2004.

[Hal 06a] M. A. K. Halliday and C. Matthiessen. *Construing experience through meaning: a language-based approach to cognition.* Continuum, 2006.

[Hal 06b] M. A. K. Halliday and J. J. Webster. *On language and linguistics.* Vol. 3, Continuum, 2006.

[Hay 94] S. Haykin. *Neural Networks: A Comprehensive Foundation.* Macmillan, New York, 1994.

[Hei 10a] A. van der Heide and G. Trivino. "A prototype of a machine speaking with emotions". In: *Actas XV Congreso Español sobre Tecnologías y Lógica Fuzzy (ESTYLF)*, pp. 55–60, 2010.

[Hei 10b] A. van der Heide and G. Trivino. "Simulating emotional personality in human computer interfaces". In: *Fuzzy Systems (FUZZ), 2010 IEEE International Conference on*, pp. 1–7, IEEE, 2010.

[How 83] G. Howard. "Frames of mind: the theory of multiple intelligences". *NY: Basics*, 1983.

[Ise 09] R. Isermann and M. Munchhof. *Identification of Dynamical Systems.* Springer-Verlag, Berlin, Heidelberg, 2009.

[Kac 00] J. Kacprzyk, R. Yager, and S. Zadroÿzny. "A fuzzy logic based approach to linguistic summaries of databases". *International Journal of Applied Mathematics and Computer Science*, No. 10, pp. 813–834, 2000.

[Kac 01] J. Kacprzyk and R. R. Yager. "Linguistic summaries of data using fuzzy logic". *International Journal of General Systems*, Vol. 30, pp. 133–154, 2001.

[Kac 08] J. Kacprzyk, A. Wilbik, and S. Zadrozny. "Linguistic summarization of time series using a fuzzy quantifier driven aggregation". *Fuzzy Sets and Systems*, Vol. 159, No. 12, pp. 1485–1499, 2008.

[Kac 10] J. Kacprzyk and S. Zadrozny. "Computing with words is an implementable paradigm: Fuzzy queries, Linguistic data summaries and Natural Language generation". *Fuzzy Systems, IEEE Transactions on*, Vol. 18, No. 3, pp. 461–472, 2010.

[Lju 98] L. Ljung. *System identification.* Springer, 1998.

[Los 04] D. E. Losada, F. Díaz-Hermida, A. Bugarín, and S. Barro. "Experiments on using fuzzy quantified sentences in adhoc retrieval". In: *Proceedings of the 2004 ACM symposium on Applied computing*, pp. 1059–1064, ACM, 2004.

[Mam 74] E. H. Mamdani. "Application of fuzzy algorithms for control of simple dynamic plant". *Proceedings of IEEE*, Vol. 121, No. 12, pp. 1585–1588, December 1974.

[Mam 75] E. H. Mamdani and S. Assilian. "An experiment in linguistic synthesis with a fuzzy logic controller". *International Journal of Man-Machine Studies*, Vol. 7, No. 1, pp. 1–13, January 1975.

[Mam 77] E. H. Mamdani. "Application of fuzzy logic to approximate reasoning using linguistic systems". *Computers, IEEE Transactions on*, Vol. 26, No. 12, pp. 1182–1191, 1977.

[Mar 12] C. Martínez-Cruz, A. van der Heide, D. Sánchez, and G. Trivino. "An approximation to the computational theory of perceptions using ontologies". *Expert Systems with Applications*, 2012.

[Men 07] J. M. Mendel. "Computing with words and its relationships with fuzzistics". *Information Sciences*, Vol. 177, No. 4, pp. 988–1006, 2007.

[Men 10] S. Mendez-Nunez and G. Trivino. "Combining Semantic Web Technologies and Computational Theory of Perceptions for Text Generation in Financial Analysis". In: *Proceedings of the 2010 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), Barcelona, Spain*, pp. 953–960, July 2010.

[Men 14] C. Menendez, L. Eciolaza, and G. Trivino. "Generating Advices with Emotional Content for Promoting Efficient Consumption of Energy". *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, Vol. 22, No. 05, pp. 677–697, 2014.

[Mor 02] J. N. Mordeson and D. S. Malik. *Fuzzy Automata and Languages: Theory and Applications (Computational Mathematics).* Chapman & Hall/CRC, March 2002.

[Nel 00] O. Nelles. *Nonlinear System Identification: From Classical Approaches to Neural Networks and Fuzzy Models.* Springer, 1 Ed., December 2000.

[Oga 67] K. Ogata. *State space analysis of control systems.* Prentice-Hall, 1967.

[Per 12] M. Pereira-Fariña, L. Eciolaza, and G. Trivino. "Quality assessment of linguistic description of data". Vol. 2, pp. 01–2012, 2012.

[Pic 00] R. W. Picard. *Affective computing.* MIT press, 2000.

[Pop 77] K. Popper and J. Eccles. *The Self and Its Brain.* Springer-Verlag, 1977.

[Rei 00] E. Reiter, R. Dale, and Z. Feng. *Building natural language generation systems.* Vol. 33, MIT Press, 2000.

[Rus 03] S. Russell and P. Norvig. *Artificial Intelligence: A Modern Approach (Second Edition).* Prentice Hall, 2003.

[Rus 69] E. H. Ruspini. "A new approach to clustering". *Information and Control*, Vol. 15, pp. 22–32, 1969.

[San 12] D. Sanchez-Valdes, L. Eciolaza, and G. Trivino. "Linguistic Description of Human Activity Based on Mobile Phone's Accelerometers". In: *Ambient Assisted Living and Home Care*, pp. 346–353, Springer, 2012.

[San 13a] D. Sanchez-Valdes, A. Alvarez-Alvarez, and G. Trivino. "Linguistic description about circular structures of the Mars surface". *Applied Soft Computing*, Vol. 13, No. 12, pp. 4738–4749, 2013.

[San 13b] D. Sanchez-Valdes and G. Trivino. "Increasing the granularity degree in linguistic descriptions of quasi-periodic phenomena". In: *Flexible Query Answering Systems*, pp. 281–292, Springer, 2013.

[San 14] D. Sanchez-Valdes and G. Trivino. "Computational Perceptions of uninterpretable data. A case study on the linguistic modeling of human gait as a quasi-periodic phenomenon". *Fuzzy Sets and Systems*, Vol. 253, pp. 101–121, 2014.

[San 15a] D. Sanchez-Valdes, A. Alvarez-Alvarez, and G. Trivino. "Dynamic linguistic descriptions of time series applied to self-track the physical activity". 2015. European Centre for Soft Computing Research Report: Mieres, Spain. *Submitted to Fuzzy Sets and Systems.*

[San 15b] D. Sanchez-Valdes, A. Alvarez-Alvarez, and G. Trivino. "Walking pattern classification using a granular linguistic analysis". 2015. European Centre for Soft Computing Research Report: Mieres, Spain. *Submitted to Applied Soft Computing.*

[San 15c] D. Sanchez-Valdes and G. Trivino. "Linguistic emotional motivation for self-tracking the physical activity". 2015. European Centre for Soft Computing Research Report: Mieres, Spain. *Submitted to Expert Systems with Applications.*

[San 68] E. S. Santos. "Maxmin automata". *Information and Control*, Vol. 13, pp. 363–377, 1968.

[Sea 97] S. R. Searle. *Linear Models.* Wiley-Interscience, 1997.

[Sod 94] T. Söderström and P. Stoica. *System identification.* Prentice-Hall, Inc, Upper Saddle River, NJ, USA, March 1994.

[Sta 09] J. H. Stapleton. *Linear Statistical Models.* Wiley, 2nd Ed., 2009.

[Tom 99] M. Tomasello. *The cultural origins of human cognition.* Harvard University Press, 1999.

[Tri 09] G. Trivino, L. Mengual, and A. Van der Heide. "Towards an architecture for semiautonomous robot telecontrol systems". *Information Sciences*, Vol. 179, No. 23, pp. 3973–3984, 2009.

[Tri 10a] G. Trivino, A. Sanchez, A. S. Montemayor, J. J. Pantrigo, R. Cabido, and E. G. Pardo. "Linguistic Description of Traffic in a Roundabout". In: *Proceedings of the 2010 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), Barcelona, Spain*, pp. 2158–2165, July 2010.

[Tri 10b] G. Trivino, A. Alvarez-Alvarez, and G. Bailador. "Application of the computational theory of perceptions to human gait pattern recognition". *Pattern Recognition*, Vol. 43, No. 7, pp. 2572–2581, July 2010.

[Tri 13] G. Trivino and M. Sugeno. "Towards linguistic descriptions of phenomena". *International Journal of Approximate Reasoning*, Vol. 54, No. 1, pp. 22–34, 2013.

[Van 11] A. Van Der Heide, D. Sánchez, and G. Trivino. "Computational models of affect and fuzzy logic". In: *Proceedings of the 7th conference of the European Society for Fuzzy Logic and Technology*, pp. 620–627, Atlantis Press, 2011.

[Yag 11] R. R. Yager. "Reasoning with doubly uncertain soft constraints". *International Journal of Approximate Reasoning*, Vol. 52, No. 4, pp. 554–561, 2011.

[Yag 82] R. R. Yager. "A new approach to the summarization of data". *Information Sciences*, Vol. 28, pp. 69–86, 1982.

[Yag 94] R. R. Yager and D. P. Filev. *Essentials of fuzzy modeling and control.* Wiley-Interscience, New York, NY, USA, June 1994.

[Yag 95] R. R. Yager. "Fuzzy Summaries in Database Mining". In: *In Proceedings 11th Conference on Artificial Intelligence for Applications*, pp. 265–269, 1995.

[Yin 02] M. Ying. "A formal model of computing with words". *Fuzzy Systems, IEEE Transactions on*, Vol. 10, No. 5, pp. 640–652, 2002.

[Zad 01] L. A. Zadeh. "A new direction in AI. Towards a Computational Theory of Perceptions of measurements to manipulation of perceptions". *AI Magazine*, Vol. 22, No. 1, 2001.

[Zad 02a] L. A. Zadeh. "Granular computing as a basis for a computational theory of perceptions". In: *Proceedings of the IEEE International Conference on Fuzzy Systems*, pp. 564–565, Honolulu, USA, May 2002.

[Zad 02b] L. A. Zadeh. "Toward a perception-based theory of probabilistic reasoning with imprecise probabilities". *Journal of Statistical Planning and Inference*, Vol. 105, pp. 233–264, 2002.

[Zad 08] L. A. Zadeh. "Toward human level machine intelligence-is it achievable? the need for a paradigm shift". *Computational Intelligence Magazine, IEEE*, Vol. 3, No. 3, pp. 11–22, 2008.

[Zad 56] L. A. Zadeh. "On the Identification Problem". *Circuit Theory, IRE Transactions on*, Vol. 3, No. 4, pp. 277–281, December 1956.

[Zad 62] L. A. Zadeh. "From circuit theory to system theory". *Proceedings of the Institute of Radio Engineers*, Vol. 50, pp. 856–865, 1962.

[Zad 65] L. A. Zadeh. "Fuzzy Sets". *Information and Control*, Vol. 8, No. 3, pp. 338–353, June 1965.

[Zad 73] L. A. Zadeh. "Outline of a new approach to the analysis of complex systems and decision processes". *Systems, Man, and Cybernetics, IEEE Transactions on*, Vol. SMC-3, No. 1, pp. 28–44, January 1973.

[Zad 75] L. A. Zadeh. "The concept of a linguistic variable and its application to approximate reasoning". *Parts I, II, and III. Information Sciences*, Vol. 8, 8, 9, pp. 199–249, 301–357, 43–80, 1975.

[Zad 83] L. A. Zadeh. "A computational approach to fuzzy quantifiers in natural languages". *Computing and Mathematics with Applications*, Vol. 9, pp. 149–184, 1983.

[Zad 96] L. A. Zadeh. "Towards a theory of fuzzy systems". In: G. J. Klir and B. Yuan, Eds., *Fuzzy Sets, Fuzzy Logic and Fuzzy Systems, Selected Papers by Lotfi Zadeh*, pp. 83–104, World Scientific Publishing Company, River Edge, NJ, USA, June 1996.

[Zad 99] L. A. Zadeh. "From computing with numbers to computing with words - from manipulation of measurements to manipulation of perceptions". *Circuits and Systems - I: Fundamental theory and applications, IEEE Transactions on*, Vol. 45(1), pp. 105–119, 1999.

Parte II

Publicaciones

Capítulo 5

Compendio de publicaciones

Este capítulo contiene el compendio de publicaciones desarrolladas durante esta tesis. Se divide en tres secciones diferentes, correspondientes a cada artículo junto con su referencia bibliográfica.

5.1. Linguistic description about circular structures of the Mars' surface

D. Sanchez-Valdes, A. Alvarez-Alvarez, and G. Trivino. "Linguistic description about circular structures of the Mars surface". *Applied Soft Computing*, Vol. 13, No. 12, pp. 4738–4749, 2013.

Applied Soft Computing 13 (2013) 4738–4749

Contents lists available at ScienceDirect

Applied Soft Computing

journal homepage: www.elsevier.com/locate/asoc

Linguistic description about circular structures of the Mars' surface

Daniel Sanchez-Valdes*, Alberto Alvarez-Alvarez, Gracian Trivino

European Centre for Soft Computing, C/ Gonzalo Gutierrez Quirós, s/n, 33600 Mieres, Asturias, Spain

A R T I C L E I N F O

Article history:
Received 2 August 2012
Received in revised form 10 June 2013
Accepted 5 August 2013
Available online 3 September 2013

Keywords:
Linguistic description of data
Computing with perceptions
Image description
Granular linguistic model of phenomena

A B S T R A C T

Satellites situated in the orbit of Mars have provided and continue providing thousand of images of the planet surface. Nevertheless the number of expert geologists analyzing these images is limited. Typically, these experts provide linguistic descriptions of their observations remarking the relevant features in the image and ignoring the irrelevant details for a given goal. In this paper, we apply our research in the field of Computational Theory of Perceptions to the challenge of developing computational systems able to generate linguistic reports comparable with the ones provided by human experts. We present a description of our contribution to solve this problem including last results of our research in this field. For example, we explore how to represent the multidimensional domain of computational perception values. We develop up the use of the relevance as an attribute of perceptions that allows us to generate reports that are automatically suited according to the user goals. We provide an application example as a demonstration of concept.

1. Introduction

Current technologies allow us to acquire and store big amounts of data about complex phenomena in many areas of science and technology. In order to be useful, these data must be interpreted and represented in an understandable way, by giving in each type of situation their relationship to data context and, in general, the information related to each specific phenomenon. Currently, these types of descriptions are reports that contain text and graphics produced by human experts. Nevertheless, the relation between the amount of data to analyze and the number of experts available is growing dramatically. This situation causes a strong demand for computational systems that can interpret and describe linguistically the large amount of information that is being generated in many areas.

This paper was motivated by the existence of a huge database of thousands of images of the Mars' surface produced by satellites and it is part of a collaboration project with the Spanish National Institute for Aerospace Technology (INTA). These images, which are usually analyzed by a small number of available experts on Martian geology, contain important information. For example, one of the most important research topics about the Mars planet is the seek of water, which can be developed by the analysis of its surface. In [1], authors use infrared to visible wavelength images of the Mars' surface to obtain relevant geological information that may help in the location of water frost. In [2], authors analyze the measurements of elevations in order to extract a highly accurate global map of the topography of Mars that determines an upper limit of the present surface water inventory.

In this work, we will use digital image processing techniques [3,4], in order to obtain automatically information about the presence of circular structures of the Mars' surface and produce a linguistic description of it. There are several examples related to this work. In [5], authors use 3D image processing to automate the surface quality analysis of exterior car body panels. From the viewpoint of describing images in natural language (NL), there are recent works based on the use of Fuzzy Logic (FL) such as [6], where authors propose a hierarchical fuzzy segmentation of the image and a collection of linguistic features able to describe each region. In [7], authors explain how a fuzzy object-relational database management system can be employed to implement and integrate the different elements needed for the linguistic description of images, namely, ontology, concept representation and language generation.

Our research is based on the Computational Theory of Perceptions (CTP) introduced by Zadeh [8–10]. In previous works in this line, we have developed computational systems able to generate linguistic descriptions of different types of phenomena. For example, we generated financial reports from data taken from the Spanish Securities Market Commission (CNMV) [11]. We have also worked in the field of activity recognition [12] and gait analysis [13], by providing automatically generated linguistic reports based on accelerometer data. And, we have also dealt with data related to the field of Intelligent Transportation Systems (ITS), where we generated assessing reports in truck driving simulators [14,15]. In [16],

* Corresponding author. Tel.: +34 985456545.
E-mail addresses: daniel.sanchezv@softcomputing.es (D. Sanchez-Valdes), alberto.alvarez@softcomputing.es (A. Alvarez-Alvarez), gracian.trivino@softcomputing.es (G. Trivino).

1568-4946/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.asoc.2013.08.003

D. Sanchez-Valdes et al. / Applied Soft Computing 13 (2013) 4738–4749 4739

we generated linguistic reports about the traffic on roundabouts and also about traffic evolution in roads [17].

In this paper, we face the challenge of creating human like reports that could help experts in Martian surface to analyze the huge database of available images.

This paper is an extension of a previous work presented in [18]. Here, we present several contributions listed as follows:

- An updated architecture of this type of computational applications, e.g., we introduce the multi-dimensionality of the domain of existence of a *Computational Perception*.
- We develop upon the concept of Relevance as a component of a *Computational Perception* that is used to choose the most suitable linguistic expressions to describe the current state of phenomena. We present a new version of *Report Generator* that is more flexible and more advanced than previous ones.
- We present an application that consists of creating linguistic reports about detected circular structures such as volcanoes or meteorite impacts. When these structures exist in the image, the report should provide information about their size and relative position according to certain criteria of the relevance of these circles for the final user.

This paper is organized as follows: Section 2 describes the architecture of a computational system able to create linguistic descriptions of phenomena. Section 3 explains how to apply this architecture for describing the presence of circular structures such as volcanoes or meteorite impacts on the Mars' surface. Section 4 shows the experimental results. And finally, Section 5 expounds some concluding remarks.

2. Linguistic description of complex phenomena

A general approach including a general architecture for computational systems for linguistic description of data can be found in [19]. We face this challenge with a more specific architecture based on our research in the field of Computational Theory of Perceptions (CTP).

CTP was introduced in Zadeh's seminal paper "From computing with numbers to computing with words – from manipulation of measurements to manipulation of perceptions" [8] and further developed in subsequent papers [9,10]. It grounds on the fact that human cognition is based on the role of perceptions, and the remarkable capability to granulate information in order to perform physical and mental tasks without any traditional measurements and computations.

In Fig. 1, we identify two main stages in the text generation process, namely, *off-line building process* and *on-line instantiation process*.

During the *off-line building process*, after analyzing the application domain and the user requirements, the designer collects a corpus of NL expressions that are typically used to describe the relevant features of the monitored phenomenon. The designer analyzes the particular meaning of each linguistic expression in specific situation types to build both, the *Granular Linguistic Model of Phenomena* (GLMP) and the *Report Template*.

During the *on-line instantiation process* the computational system uses input data obtained by the *Data Acquisition Module* to generate suitable linguistic reports as result of two instantiation processes of these two generic data structures that are performed by the *Validity Module* and the *Expression Module* respectively.

2.1. DAQ Module

The *Data acquisition (DAQ) module* consists of sensors and measurement hardware-software resources. In our approach, the *DAQ Module* provides the data needed to feed up the instantiation process. It provides the interface with the application physical environment. This module could include either sensors or access to information stored in a database. In Section 3.1, we explain how the *DAQ Module* takes the information from satellite images and implements an image processing algorithm.

2.2. Granular linguistic model of a phenomenon

The GLMP data structure is a useful paradigm for developing computational systems able to generate linguistic descriptions of data [13,17,18]. The main element of this structure is known as Computational Perception (CP), which is based on the concept of linguistic variable developed by Zadeh [20–22]. CPs are computational models of information units (granules) acquired by the designer about the phenomenon to be modeled, e.g., CPs correspond to perceptions of specific parts of the phenomenon at certain granularity degree.

The GLMP is a network of Perception Mappings (PM), which are the elements used to create and aggregate or combine CPs. Each PM receives a set of input CPs and transmits upwards an output CP. We say that each output CP is explained by the PM using a set of input CPs. In this network, each CP covers specific aspects of the phenomenon with certain granularity degree. We call first order perception mappings (1PM) to those PMs which serve as input to the GLMP. We call first order computational perceptions (1CP) to the output of a 1PM. The PMs whose input are CPs are called 2PMs and their outputs are 2CPs. This classification is inspired on the definition of the three worlds by Popper, namely, world-1 of physical objects (phenomena), world-2 of the perceived objects (1CP) and world-3 of the mental objects built by using the objects in the world-2 (2CP) [23].

Fig. 2 shows an example of a GLMP. In this example, the phenomenon can be described at a very basic level in terms of three variables providing values z_1, z_2, and z_3 respectively. In this GLMP, other two higher-level descriptions of the phenomenon are provided. These descriptions are given in the form of computational perceptions $2CP_4$ and $2CP_5$. The second order perception mappings (2PMs) $2PM_4$ and $2PM_5$ indicate that $2CP_4$ and $2CP_5$ can be explained in terms of $1CP_1$, $1CP_2$, and $1CP_3$. Finally, the top-order

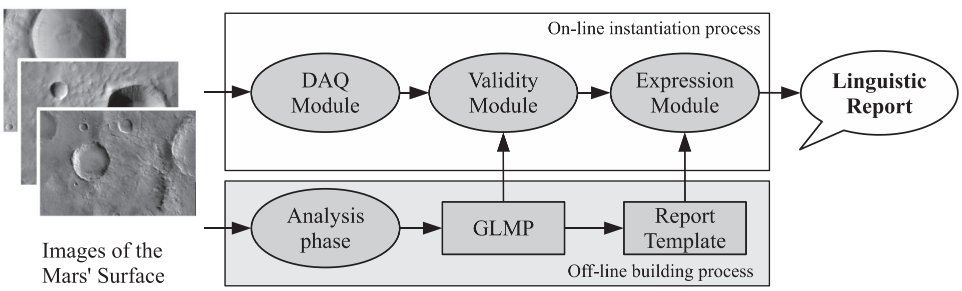


Fig. 1. Architecture of a computational system for generating linguistic description of data.

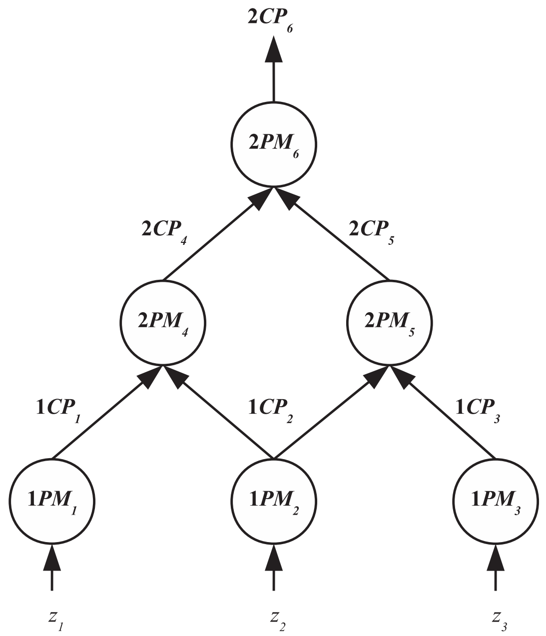


Fig. 2. Example of a GLMP.

description of the phenomenon is provided, at the highest level of abstraction, by $2CP_6$, explained by $2PM_6$ in terms of $2CP_4$ and $2CP_5$. Notice that, by using this structure, one can provide not only a linguistic description of the phenomenon at a certain level, but an explanation in terms of linguistic expressions at a lower level. In the following subsections, we will see how, after being instantiated with a set of input data, the GLMP provides a structure of valid sentences that in medium size applications could include hundreds of sentences.

2.2.1. Computational perception

A CP is a tuple with three components (A, W, R) described as follows:

A is a multidimensional matrix of linguistic expressions (words or sentences in NL) that represents the whole linguistic domain of CP, i.e., all its possible linguistic values. In the application context, each element of A describes the value of CP in each situation of the phenomenon with specific granularity degree. During the *off-line stage*, these values are defined by the designer extracting suitable sentences from the linguistic corpus of the application domain. For example, consider the sentences which describe a CP that were generated with the template:

"The image contains {*zero* | *one* | *two* | *three* | *four* | *various* | *many*} {*small* | *medium* | *big*} *circles"* Here, because the linguistic template contains two variables, we will organize the domain of A as a bidimensional matrix:

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{41} & a_{42} & a_{43} \\ a_{51} & a_{52} & a_{53} \\ a_{61} & a_{62} & a_{63} \\ a_{71} & a_{72} & a_{73} \end{pmatrix}$$

Note that the number of dimensions and size of A depends on each application domain. In this case, rows refer to number of circles (7 rows: *zero, one, two, . . ., many*) and columns to the size of circles (3 columns: *small, medium, big*).

W is a multidimensional matrix of validity degrees whose elements are in the interval [0, 1] assigned to each element of A in a specific context. The validity degree represents the precision of each sentence when describing the specific input data. During the *on-line stage*, these validity values are assigned dynamically reflecting the current state of the monitored phenomenon. For example, provided the NL expressions in A, the matrix of validity degrees will be also a bidimensional matrix:

$$W = \begin{pmatrix} w_{11} & w_{12} & w_{13} \\ w_{21} & w_{22} & w_{23} \\ w_{31} & w_{32} & w_{33} \\ w_{41} & w_{42} & w_{43} \\ w_{51} & w_{52} & w_{53} \\ w_{61} & w_{62} & w_{63} \\ w_{71} & w_{72} & w_{73} \end{pmatrix}$$

R is a multidimensional matrix of relevancy degrees whose elements are also in the interval [0, 1] assigned to each element of A in a specific context. The value of relevance depends on the application. During the *off-line stage*, these values are assigned by the designer depending on the specific application requirements. The designer assigns values to R indicating the relative importance of the described situation types. During the report generation phase, the values of R will be used as complementary data to select the most suitable sentences to describe the current state of the phenomenon. For example, provided the NL expressions in A, the matrix of relevancy degrees will be also a bidimensional matrix:

$$R = \begin{pmatrix} 0 & 0 & 0.2 \\ 0 & 0 & 0.6 \\ 0 & 0 & 0.7 \\ 0 & 0 & 0.8 \\ 0 & 0 & 0.9 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

Here, the designer has defined that, the bigger and numerous the circles, the more relevant is the linguistic expression; while the small and medium circles are not relevant at all.

2.2.2. Perception mapping

A PM is a tuple (U, y, g, T) where:

U is a set of input CPs. As explained before, we call first order perception mappings (1PMs) when U are values $z \in \mathbb{R}$ provided either by sensors or obtained from a database.

y is the output CP, $y = (A_y, W_y, R_y)$.

g is an aggregation function that calculates the vector of validity degrees W_y. It is a fuzzy aggregation of the validity degrees of the input CPs. In FL, many different types of aggregation functions can be developed. For example g, could be implemented using a set of fuzzy rules or, in the case of 1PMs, g is built using a set of membership functions as follows:

$$W_y = \begin{pmatrix} w_{11} & w_{12} & \cdots & w_{1n} \\ w_{21} & w_{22} & \cdots & w_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ w_{m1} & w_{m2} & \cdots & w_{mn_y} \end{pmatrix} = \begin{pmatrix} \mu_{a_{11}}(z) & \mu_{a_{12}}(z) & \cdots & \mu_{a_{1n}}(z) \\ \mu_{a_{21}}(z) & \mu_{a_{22}}(z) & \cdots & \mu_{a_{2n}}(z) \\ \vdots & \vdots & \vdots & \vdots \\ \mu_{a_{m1}}(z) & \mu_{a_{m2}}(z) & \cdots & \mu_{a_{mn_y}}(z) \end{pmatrix}$$

where W_y is the matrix of validity degrees associated to the matrix of linguistic expressions A_y, and $\mu_{a_{ij}}(z)$ is the membership degree of the input data to each membership function.

T is a text generation algorithm that allows generating the sentences in A_y. For example, in simple cases, T is a linguistic template, e.g., *"The circle is* $\{$*small* | *medium* | *big*$\}$*"*.

2.3. Validity Module

This processing module calculates the validity degrees of each CP during the *on-line process* by instantiating in the GLMP structure the most suitable linguistic expressions to represent the input data, i.e., this module provides as output a collection of linguistic clauses together with associated validity degrees.

2.4. Report template

The Report Template is designed during the *off-line process*. Here, we develop upon the concept of Fuzzy Tree of Choices (FTC), which is a mechanism to represent part of the constraints imposed to the linguistic report. It consists of a directed graph including choices and the linguistic expressions to be linked together. This FTC has a generic version that will be instantiated for each input data (see Fig. 7).

2.5. Expression Module

Provided a set of valid linguistic clauses, this processing module is in charge of combining this information to generate the most relevant linguistic report by choosing and connecting the adequate ones. This is done during the *on-line process* by instantiating the *Report template*.

The *Expression Module* is used to analyze every combination of branches in the FTC until it finds out the sequence that accumulates the highest validity degree. More specifically, we use Eq. (1) to calculate the product of the validity degrees of all linguistic expressions belonging to each possible sequence.

$$w_{\text{sequence}} = \prod_{\forall i \in \text{sequence}} w_i \tag{1}$$

Once we obtain the sequence with the highest validity degree, the relevance of each expression is used to define the order in which the information is going to be expressed in the final report. In the same way, it determines which information is not relevant to be included in such report. So, the FTC is pruned and reordered, resulting in an instantiated report depending on the analyzed image.

3. Linguistic description about the circular structures

In this section, we describe in detail a computational application for producing linguistic descriptions of the Mars' surface.

3.1. DAQ Module: image processing

This module is in charge of recognizing circles in the image. The recognition of patterns is an open issue of research in the field of automatic image processing. Here, in order to extract information about the presence of circular structures, we have applied classical filtering techniques to remove the background, techniques of edge detection, and the generalized Hough transform which is particularly suitable for detecting the presence of circles [4,24].

In the first task, known as pre-processing, the image is converted to grayscale and everything that is not interesting in the image is "deleted". This is a procedure that slightly transforms the image to eliminate any noise, imperfections, shine, etc. Then, we detect the edges in the image that can be defined as transitions between two significantly different levels of gray intensity. This provides valuable information on the borders of objects that can be used for image segmentation.

We can find some interesting edge detection and segmentation methods based on Soft Computing techniques. For example, an ant-inspired algorithm was used in [25] for the detection of image edge features while a simulated annealing was used in [26] to edge-based segmentation of grayscale images. There are also recent works that use fuzzy techniques for the same purpose such as [27], where authors used fuzzy c-means clustering with weighted image patch for image segmentation; in [28], a color image segmentation was practiced using fuzzy clustering techniques and competitive neural networks; and, finally, authors in [29], used an adaptive region based color texture segmentation using a fuzzified distance metric. Here, in order to solve the faced problems in this prototype, we have used a classical version of the Sobel operator [4] as

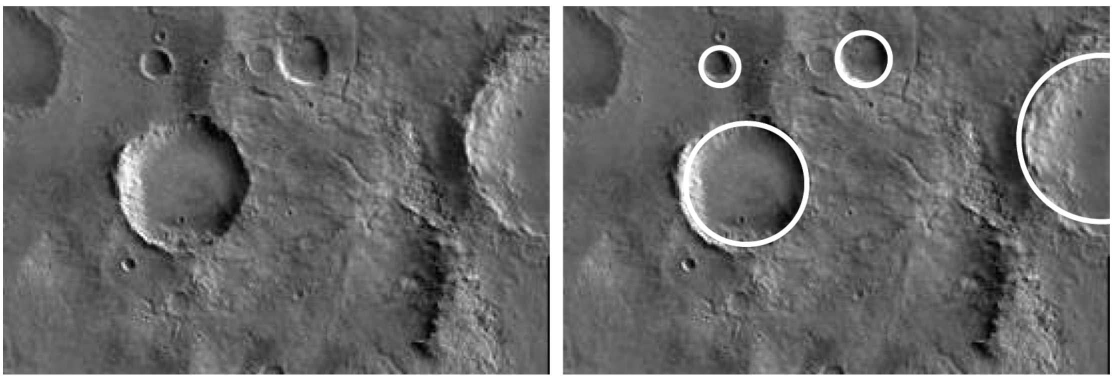


Fig. 3. Example of circles detection in an image.

boundary filter which consists of two arrays whose size is 5×5 pixels.

The search for circles in the image is tackled by means of the Hough Transform [30], which was initially concerned with the identification of lines in the image, but later it has been extended to identify arbitrary shapes such as circles or ellipses. To get good results with this approach, the pre-processing stage is essential, since the procedure is mainly based on the color jump produced at the edge of either meteorite impacts or volcano craters. The great difficulty of this procedure lies in the analysis of those images in which, by its nature, the land has many irregularities. This fact accumulates inaccuracies in the analysis and, therefore, we must assume a margin of error in the obtained results. Moreover, since we do not know in advance the size and shapes of the relevant objects, we have to work with a generalized Hough transform that enables the detection of objects whose shapes and dimensions are, initially, unknown. As any other pattern recognition algorithm, the Hough transform introduces some uncertainty by committing several false positives and false negatives mistakes. We have taken advantage of this feature to emulate the behavior of a human observer that could make similar mistakes and therefore to generate linguistic descriptions containing such type of imprecision.

For each image, the *DAQ Module* provides the number of detected circles in the image, together with their x and y centroid coordinates, and their corresponding radius (in pixels). As an example of the performance of this module, Fig. 3 shows the circles detected for a specific image.

3.2. Validity Module

Fig. 4 shows the designed GLMP. We have defined three 1CPs which describe the size and position (in the vertical and horizontal axis respectively) of each circle and five 2CPs which describe the quantity of circles and their position at different levels of detail. The 2PMs that appear in this GLMP can be classified in those ones that aggregate (Σ) the information from the same subordinate CP and the ones which combine (Π) information from different subordinate CPs. In the following subsections, we thoroughly describe each PM that appears in the GLMP.

3.2.1. Position of a circle in the x coordinate ($1PM_X$)

The elements of this 1PM can be described as follows:

U is the numerical value ($z_1 \in \mathbb{R}$) of the x coordinate of the centroid of a circle.

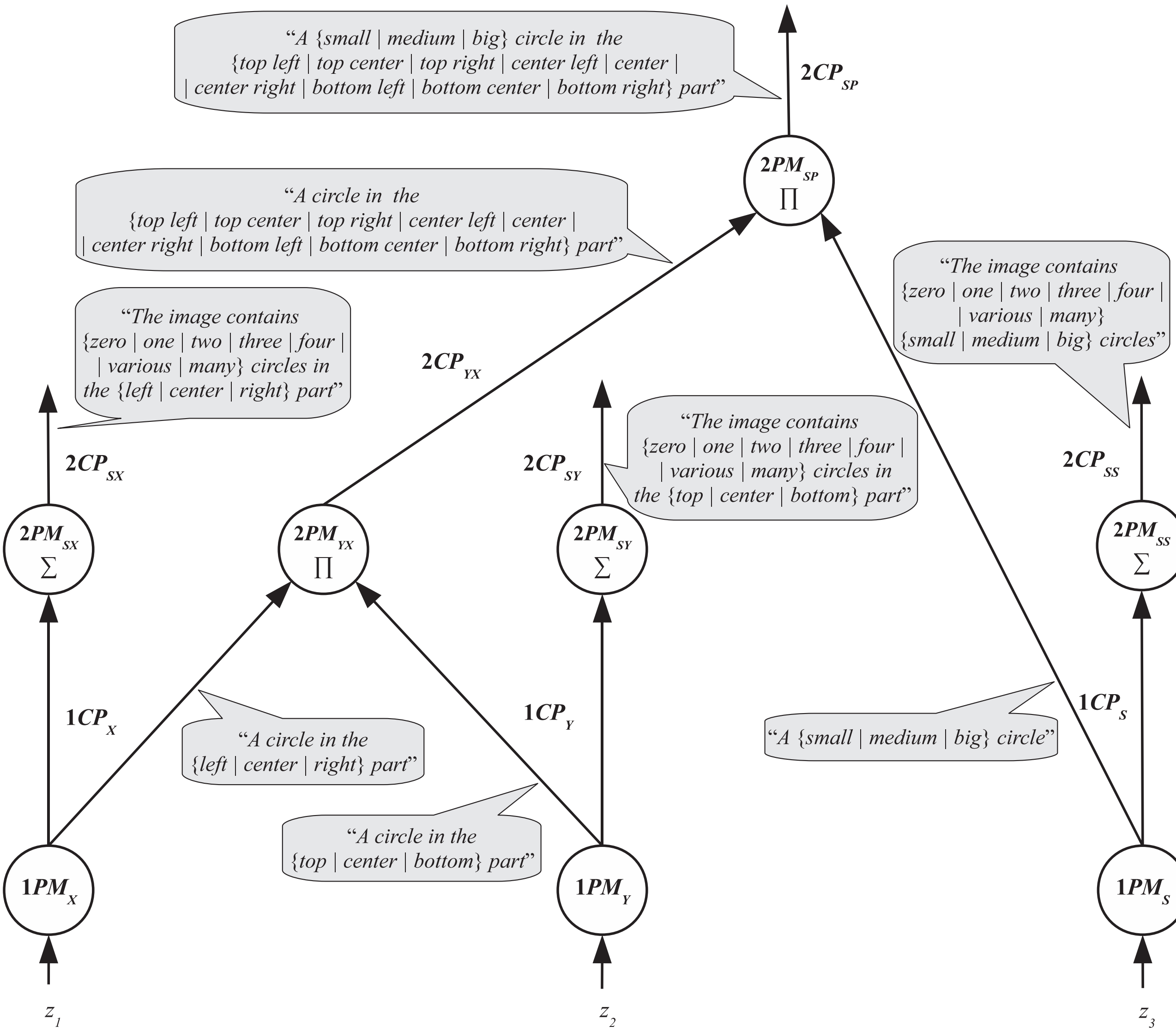


Fig. 4. GLMP for the linguistic description of the Mars' surface.

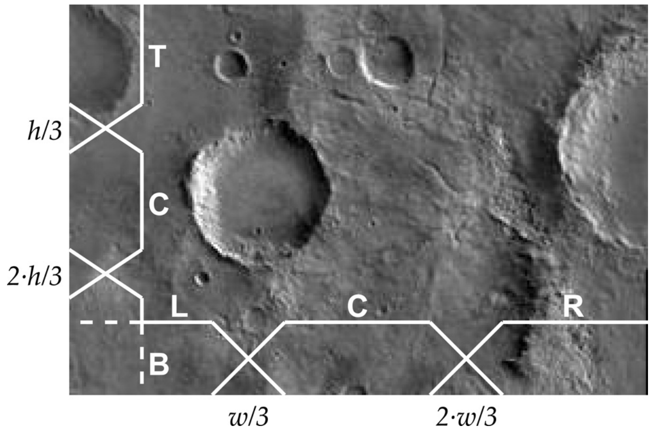


Fig. 5. Linguistic labels that represent the position attending to the x and y coordinates for a specific image.

y is the output CP ($1CP_X$), that has the following set of possible sentences a_{X_i}, with $i \in \{1, 2, 3\}$: *"A circle in the* $\{$*left* | *center* | *right*$\}$ *part"*, e.g., a_{X_1} = *"A circle in the left part"*. Since this CP works only as input data and it is not used in the linguistic reports, the relevance values r_{X_i} are set to zero $\forall i$.

g is the function that calculates the validity degrees of the output CP. These values are obtained by means of uniformly distributed trapezoidal membership functions (MFs) forming a strong fuzzy partition (SFP) [31]. Here, the three linguistic labels associated to each NL expression are trapezoidal MFs distributed over one and two thirds of the image width w. We have heuristically decided its vertexes as follows: $\{left \rightarrow (-\infty, -\infty, 7 \cdot w/24, 9 \cdot w/24)$, $center \rightarrow (7 \cdot w/24, 9 \cdot w/24, 15 \cdot w/24, 17 \cdot w/24)$, $right \rightarrow (15 \cdot w/24, 17 \cdot w/24, \infty, \infty)\}$ in such a way that these vertexes are placed in one and two thirds of the image width plus or minus an additional value corresponding to $w/24$. In Fig. 5, can be seen a graphical representation of these linguistic labels for a specific image.

3.2.2. Position of a circle in the y coordinate ($1PM_Y$)

This 1PM is similar to $1PM_X$ and can be described as follows:

U is the numerical value ($z_2 \in \mathbb{R}$) of the y coordinate of the centroid of a circle.

y is the output CP ($1CP_Y$), it has the following set of possible sentences a_{Y_i}, with $i \in \{1, 2, 3\}$: *"A circle in the* $\{$*top* | *center* | *bottom*$\}$ *part"*. Since this CP also works only as input data and it is not used in the linguistic reports, the relevance values r_{Y_i} are zero $\forall i$.

g is the output function. It consists of three trapezoidal MFs that were distributed over one and two thirds of the image height h similarly to the position of the circle in the x coordinate, resulting in the following vertexes: $\{top \rightarrow (-\infty, -\infty, 7 \cdot h/24, 9 \cdot h/24)$, $center \rightarrow (7 \cdot h/24, 9 \cdot h/24, 15 \cdot h/24, 17 \cdot h/24)$, $bottom \rightarrow (15 \cdot h/24, 17 \cdot h/24, \infty, \infty)\}$. In Fig. 5, can be seen a graphical representation of these linguistic labels for a specific image.

3.2.3. Size of a circle ($1PM_S$)

This 1PM can be described as follows:

U is the numerical value ($z_3 \in \mathbb{R}$) of the radius of the circle in pixels.

y is the output CP ($1CP_S$), it has the following set of possible sentences a_{S_i}, with $i \in \{1, 2, 3\}$: *"A*$\{$*small* | *medium* | *big*$\}$*circle"*. Since this CP also works only as input data and it is not used in the linguistic reports, the relevance values r_{S_i} are zero $\forall i$.

g is the output function. It consists of three trapezoidal linguistic labels heuristically defined by their vertexes as follow: $\{$*small* (S) $(-\infty, -\infty, s/16, s/8)$, *medium* (M) $(s/16, s/8, 2 \cdot s/3, 3 \cdots /2)$, *big* (B) $(2 \cdot s/3, 3 \cdot s/2, \infty, \infty)\}$, where s is the minimum between the width (w) and height (h) of the figure divided by six, as is shown in Eq. (2):

$$s = \frac{\min(w, h)}{6}. \tag{2}$$

3.2.4. Summary of circles in the x position ($2PM_{SX}$)

This 2PM aggregates (Σ) the information from the $1CP_X$ subordinate CP. It has the following elements:

U is the input CP: $1CP_X$.

y is the output CP ($2CP_{SX}$), it has the following set of possible sentences $a_{SX_{ij}}$, with $i \in \{1, 2, \ldots, 7\}$ and $j \in \{1, 2, 3\}$:

$a_{SX_{1j}} \rightarrow$ *"The image does not contain any circle in the* $\{$*left* | *center* | *right*$\}$ *part"*

$a_{SX_{2j}} \rightarrow$ *"The image contains one circle in the* $\{$*left* | *center* | *right*$\}$ *part"*

$a_{SX_{3j}} \rightarrow$ *"The image contains two circles in the* $\{$*left* | *center* | *right*$\}$ *part"*

$a_{SX_{4j}} \rightarrow$ *"The image contains three circles in the* $\{$*left* | *center* | *right*$\}$ *part"*

$a_{SX_{5j}} \rightarrow$ *"The image contains four circles in the* $\{$*left* | *center* | *right*$\}$ *part"*

$a_{SX_{6j}} \rightarrow$ *"The image contains various circles in the* $\{$*left* | *center* | *right*$\}$ *part"*

$a_{SX_{7j}} \rightarrow$ *"The image contains many circles in the* $\{$*left* | *center* | *right*$\}$ *part"*

In this specific application, we consider that this information is not relevant in the final report but, it is available if the designer decides to show it in the future. Therefore, the relevance values $r_{SX_{ij}}$ are zero $\forall i, j$.

g is the output function, which is based on the α-cuts based method proposed in [32]. For each part of the image that appears in the set of sentences (*left*, *center*, or *right*) determined by the index j, we calculate the percentage of circles contained at each α-level ($N_{\alpha j}$) by means of Eq. (3), with $\alpha \in A = \{0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1\}$:

$$N_{\alpha j} = \frac{1}{K} \sum_{k=1}^{K} F_\alpha(w_{X_{1j}}[k]) \tag{3}$$

where $F_\alpha(z)$ is defined by Eq. (4), and $w_{X_{1j}}[k]$ is the validity degree of $1CP_X$, which is calculated for each circle (k) until the total number of circles (K) detected in the image.

$$F_\alpha(z) = \begin{cases} 1 & \text{if} \quad z \geq \alpha \\ 0 & \text{if} \quad z < \alpha \end{cases} \tag{4}$$

Then, we calculate the membership degree of each $N_{\alpha j}$ to each element of the set of linguistic quantifiers: $\{Q_1, \ldots, Q_7\} = \{Zero, One, Two, Three, Four, Various, Many\}$, e.g., $\mu_{Q_4}(N_{\alpha j}) = Three(N_{\alpha j})$. The shapes of these linguistic labels are determined by the total number of circles K as can be seen in Fig. 6.

The last step is to calculate the average value of the membership degrees obtained for each α-level using Eq. (5). The number of elements in the set A is the resolution degree, i.e., here, $|A| = 10$.

$$w_{SX_{ij}} = \frac{1}{|A|} \sum_{\forall \alpha \in A} \mu_{Q_i}(N_{\alpha j}) \tag{5}$$

This final value contains the relevant information about the amount of circles attending to the x coordinate (left, center or right) of their centroid. For example, the validity degree of the sentence

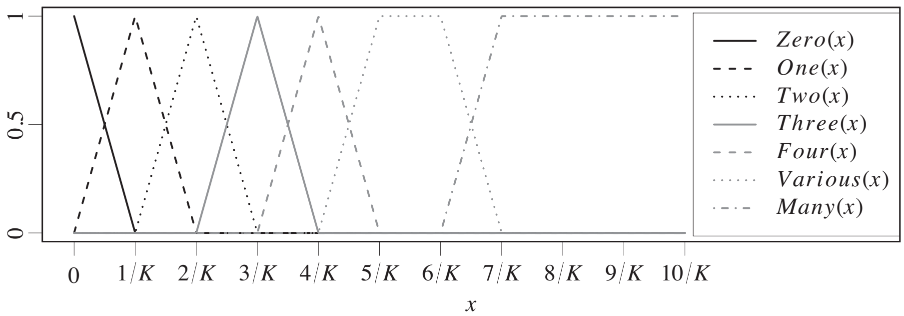


Fig. 6. Linguistic labels that represent the linguistic quantifiers "*Zero*", "*One*", "*Two*", "*Three*", "*Four*", "*Various*", or "*Many*" circles.

"The image contains three circles in the center part" ($w_{SX_{42}}$) will be determined by Eq. (6):

$$w_{SX_{42}} = \frac{1}{|\mathrm{A}|}\sum_{\forall\alpha\in\mathrm{A}} Three(N_{\alpha 2}) \tag{6}$$

3.2.5. Summary of circles in the y position ($2PM_{SY}$)

This 2PM aggregates (Σ) the information from the $1CP_Y$ subordinate CP in the same way as $2PM_{SX}$. It has the following elements:

U is the input CP: $1CP_Y$.

y is the output CP ($2CP_{SY}$), it has the following set of possible sentences $a_{SY_{ij}}$, with $i \in \{1, 2, \ldots, 7\}$ and $j \in \{1, 2, 3\}$:

$a_{SY_{1j}} \rightarrow$ *"The image does not contain any circle in the* {*top* | *center* | *bottom*} *part"*
$a_{SY_{2j}} \rightarrow$ *"The image contains one circle in the* {*top* | *center* | *bottom*} *part"*
$a_{SY_{3j}} \rightarrow$ *"The image contains two circles in the* {*top* | *center* | *bottom*} *part"*
$a_{SY_{4j}} \rightarrow$ *"The image contains three circles in the* {*top* | *center* | *bottom*} *part"*
$a_{SY_{5j}} \rightarrow$ *"The image contains four circles in the* {*top* | *center* | *bottom*} *part"*
$a_{SY_{6j}} \rightarrow$ *"The image contains various circles in the* {*top* | *center* | *bottom*} *part"*
$a_{SY_{7j}} \rightarrow$ *"The image contains many circles in the* {*top* | *center* | *bottom*} *part"*

Similarly to $2PM_{SX}$, we consider that this information is not relevant in the final report, but it is available if the expert decides to show it in the future. Therefore, the relevance values $r_{SY_{ij}}$ are zero $\forall i, j$.

g is the output function, which calculates the validity degrees ($w_{SY_{ij}}$) in the same way as the validity degrees of $2CP_{SX}$ ($w_{SX_{ij}}$), e.g., the validity degree of the sentence *"The image contains four circles in the top part"* ($w_{SY_{51}}$) will be determined by Eq. (7):

$$w_{SY_{51}} = \frac{1}{|\mathrm{A}|}\sum_{\forall\alpha\in\mathrm{A}} Four\left(\frac{1}{K}\sum_{k=1}^{K} F_\alpha(w_{Y_{11}}[k])\right) \tag{7}$$

3.2.6. Summary of circles attending to the size ($2PM_{SS}$)

This 2PM aggregates (Σ) the information from the $1CP_S$ subordinate CP in the same way as $2PM_{SX}$ and $2PM_{SY}$. It has the following elements:

U is the input CP: $1CP_S$.

y is the output CP ($2CP_{SS}$), it has the following set of possible sentences $a_{SS_{ij}}$, with $i \in \{1, 2, \ldots, 7\}$ and $j \in \{1, 2, 3\}$:

$a_{SS_{1j}} \rightarrow$ *"The image does not contain any* {*small* | *medium* | *big*} *circle"*
$a_{SS_{2j}} \rightarrow$ *"The image contains one* {*small* | *medium* | *big*} *circle"*
$a_{SS_{3j}} \rightarrow$ *"The image contains two* {*small* | *medium* | *big*} *circles"*
$a_{SS_{4j}} \rightarrow$ *"The image contains three* {*small* | *medium* | *big*} *circles"*
$a_{SS_{5j}} \rightarrow$ *"The image contains four* {*small* | *medium* | *big*} *circles"*
$a_{SS_{6j}} \rightarrow$ *"The image contains various* {*small* | *medium* | *big*} *circles"*
$a_{SS_{7j}} \rightarrow$ *"The image contains many* {*small* | *medium* | *big*} *circles"*

In this application, the designer decided that this 2CP is very relevant for the report. On the one hand, the relevancy matrix shows as the absence of small and medium circles is not relevant for final report ($r_{SS_{11}} = r_{SS_{12}} = 0$). On the other hand, the presence of various or many big circles is the most relevant information ($r_{SS_{63}} = r_{SS_{73}} = 1$), meanwhile the existence of medium circles is somehow relevant:

$$R_{SS} = \begin{pmatrix} 0 & 0 & 1 \\ 0.4 & 0.6 & 0.8 \\ 0.4 & 0.6 & 0.8 \\ 0.4 & 0.6 & 0.8 \\ 0.4 & 0.6 & 0.8 \\ 0.8 & 0.9 & 1 \\ 0.8 & 0.9 & 1 \end{pmatrix}$$

g is the output function, which calculates the validity degrees ($w_{SS_{ij}}$) in the same way as the validity degrees of $2CP_{SX}$ and $2CP_{SY}$, e.g., the validity degree of the sentence *"The image contains many big circles"* ($w_{SS_{73}}$) will be determined by Eq. (8):

$$w_{SS_{73}} = \frac{1}{|\mathrm{A}|}\sum_{\forall\alpha\in\mathrm{A}} Many\left(\frac{1}{K}\sum_{k=1}^{K} F_\alpha(w_{S_{13}}[k])\right) \tag{8}$$

3.2.7. Position of a circle in y and x coordinates ($2PM_{YX}$)

This 2PM combines (Π) the information from $1CP_Y$ and $1CP_X$ subordinate CPs. It has the following elements:

U are the input CPs: $\{1CP_Y, 1CP_X\}$

y is the output CP ($2CP_{YX}$), it has the following set of possible sentences a_{YX_i}, with $i \in \{1, 2, \ldots, 9\}$: *"A circle in the* {*top left* | *top center* | *top right* | *center left* | *center* | *center rigth* | *bottom left* | *bottom center* | *bottom right*} *part"*. We consider that this information is not relevant in the final report, but it is available if the designer decides to show it in the future. Therefore, the relevance values r_{YX_i} are set to zero $\forall i$.

g is the output function, which calculates the matrix of validity degrees (W_{YX_i}), using the product for the intersection, as follows:

$$W_{YX} = \begin{pmatrix} w_{Y_1}\cdot w_{X_1} & w_{Y_1}\cdot w_{X_2} & w_{Y_1}\cdot w_{X_3} & w_{Y_2}\cdot w_{X_1} & \ldots & w_{Y_3}\cdot w_{X_3} \end{pmatrix}$$

D. Sanchez-Valdes et al. / Applied Soft Computing 13 (2013) 4738–4749 4745

3.2.8. Size of a circle and its position x, y ($2PM_{SP}$)

This 2PM combines (Π) the information from $2CP_{YX}$ and $1CP_S$ subordinate CPs. It has the following elements:

U are the input CPs: $\{1CP_{YX}, 1CP_S\}$

y is the output CP ($2CP_{SP}$), it has the following set of possible sentences $a_{SP_{ij}}$, with $i \in \{1, 2, 3\}$ and $j \in \{1, \ldots, 9\}$:

$a_{SP_{1j}} \rightarrow$ *"A small circle in the* $\{$*top left* | *top center* | *top right* | *center left* | *center* | *center right* | *bottom left* | *bottom center* | *bottom right*$\}$ *part"*

$a_{SP_{2j}} \rightarrow$ *"A medium circle in the* $\{$*top left* | *top center* | *top right* | *center left* | *center* | *center right* | *bottom left* | *bottom center* | *bottom right*$\}$ *part"*

$a_{SP_{3j}} \rightarrow$ *"A big circle in the* $\{$*top left* | *top center* | *top right* | *center left* | *center* | *center right* | *bottom left* | *bottom center* | *bottom right*$\}$ *part"*

In this case, we are not too much interested in small circles position. We want to emphasize the presence of medium and, specially, big circles in the image. Therefore, the set of relevance values is the following:

$$R_{SP} = \begin{pmatrix} 0.4 & 0.4 & 0.4 & 0.4 & 0.4 & 0.4 & 0.4 & 0.4 & 0.4 \\ 0.7 & 0.7 & 0.7 & 0.7 & 0.7 & 0.7 & 0.7 & 0.7 & 0.7 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}$$

g is the output function, which calculates the validity degrees ($w_{SP_{ij}}$) based on the validity degrees of the subordinate CPs, using for the intersection the product of the two matrices of validity degrees, as follows:

$$W_{SP} = \begin{pmatrix} w_{S_1} \\ w_{S_2} \\ w_{S_3} \end{pmatrix} \cdot \begin{pmatrix} w_{YX_1} & w_{YX_2} & w_{YX_3} & w_{YX_4} & w_{YX_5} & w_{YX_6} & w_{YX_7} & w_{YX_8} & w_{YX_9} \end{pmatrix}$$

3.3. Expression Module

In this application, we developed the FTC that can be seen in Fig. 7. It contains four choices (queries) made over $2CP_{SS}$ listed as follows:

$q_1 \rightarrow$ Does the image contain circles?
$q_2 \rightarrow$ How many big circles does the image contain?
$q_3 \rightarrow$ How many medium circles does the image contain?
$q_4 \rightarrow$ How many small circles does the image contain?

As we explained in Section 2.5, the relevancy matrix of each CP (R) feed the *Expression Module* to instantiate the FTC. Therefore, each report will be different depending on the analyzed image. In

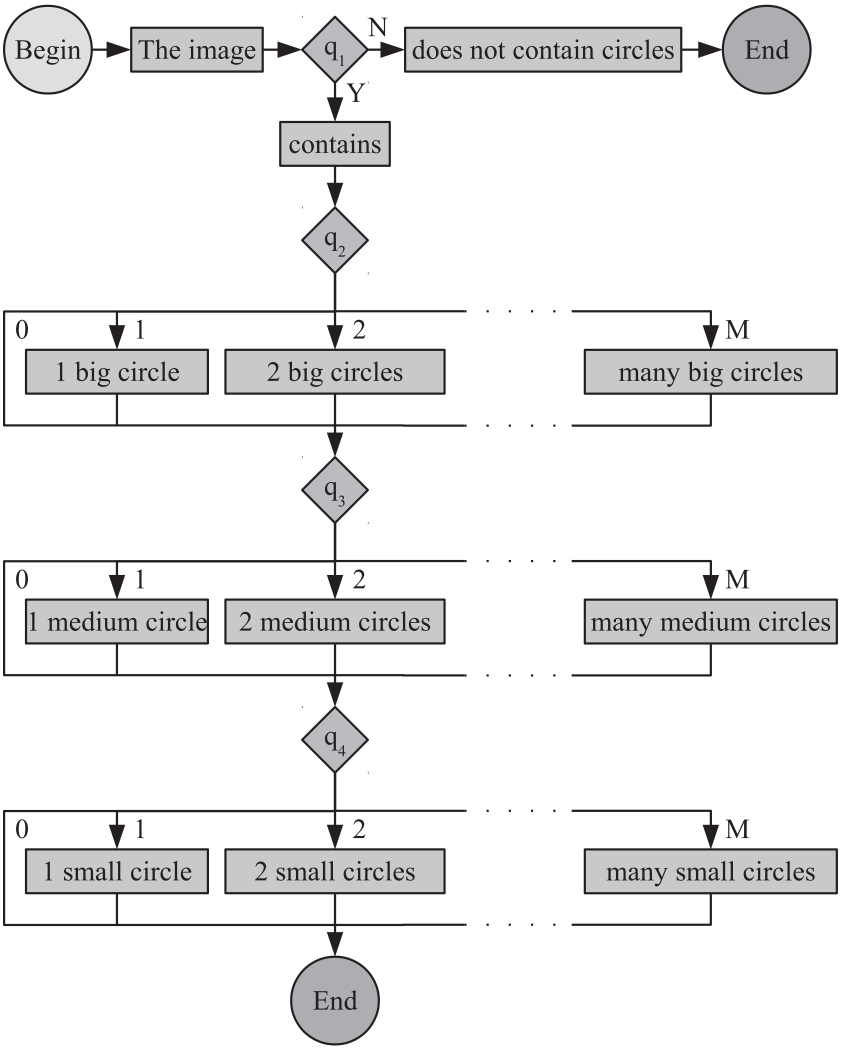


Fig. 7. Report template. Fuzzy tree of choices.

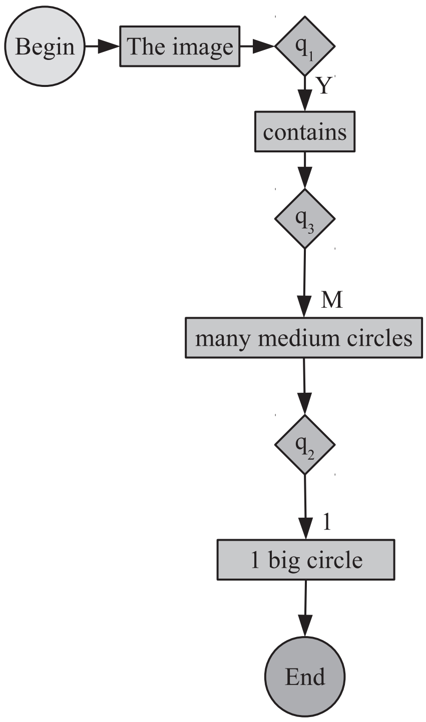


Fig. 8. Report automatically generated for the first example.

this application example, the elements of the relevancy matrix of all CPs are zero, except for $2CP_{SS}$ (R_{SS}) and $2CP_{SP}$ (R_{SP}). This is because the designer only wants to include in the report the information corresponding to these two CPs.

The $2CP_{SS}$ informs about the number of circles in the image. R_{SS} specifies the relevancy values of each possible sentence to be included in the final report. Figs. 8 and 9 show two examples of reports automatically generated by the *Expression Module* for two different images (Images 1 and 5 of the experimentation).

The $2CP_{SS}$ of the first example, Fig. 8, gave as result that the image contains four small circles, many medium circles and one big circle, as the sequence with highest validity degree calculated by Eq. (1). Attending to the relevancy matrix R_{SS}, the relevancy degrees corresponding to this result are 0.4, 0.9 and 0.8, respectively. These results allow the *Expression Module* to instantiate the corresponding report.

As part of the *report template*, we have introduced a relevancy threshold (T_{SS}) associated with the expression of CP_{SS}. It indicates the minimum relevancy degree of sentences to be included in the report. In this first example, with T_{SS} fixed to 0.5, the final report will include information about medium and big circles, but not about small ones. Therefore, the resulting sentence from this example would be *"The image contains many medium circles and one big circle"*.

The $2CP_{SS}$ of the second example, Fig. 9, gave as result that the analyzed image contains various small circles, three medium circles and one big circle. As we can see in this figure, the report created by the *Expression Module* in this case is totally different from the previous one. If we look at R_{SS}, the relevancy degrees corresponding to this result are equal to 0.8, 0.6 and 0.8, respectively. Since the relevance of having various small circles is the same as having one big circle, we established that the presence of big circles has preference in the final report. With T_{SS} = 0.5, the final report will include the following sentence *"The image contains one big circle, various small circles and three medium circles"*. Note that, the quantity of information provided and the order in which this information has been expressed has completely changed.

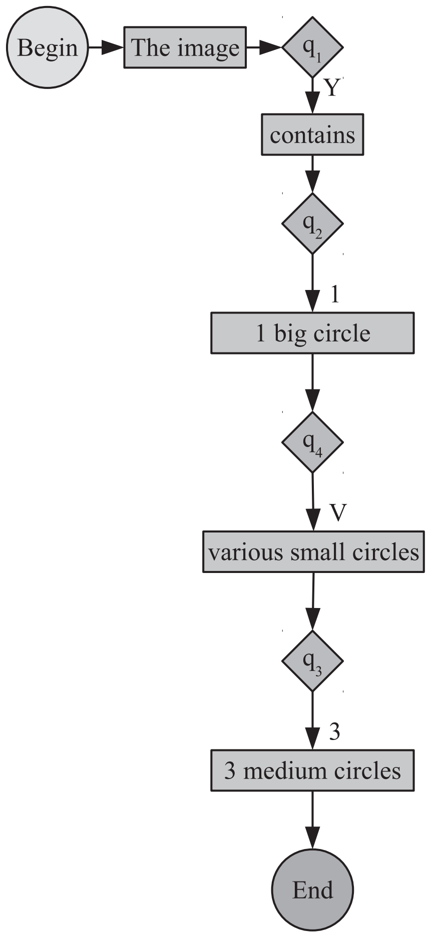


Fig. 9. Report automatically generated for the second example.

Once the relevant number of circles is described, the designer would like to include their positions in the image. This information is provided by the $2CP_{SP}$. The relevancy degrees of this CP are specified in R_{SP}. In this case, we also define a relevancy threshold (T_{SP}) that allows the *Expression Module* to know the locations of which circles have to be included in the final report. Following with the examples showed in Figs. 8 and 9, with T_{SP} fixed to 0.8, the final report of the first example will add that *"the big circle is in the center of the image"* and the final report of the second example will add that *"the big circle is in the center left part of the image"*. In both cases, the information about the location of small and medium circles has been omitted (their relevancy degrees are lower than T_{SP}).

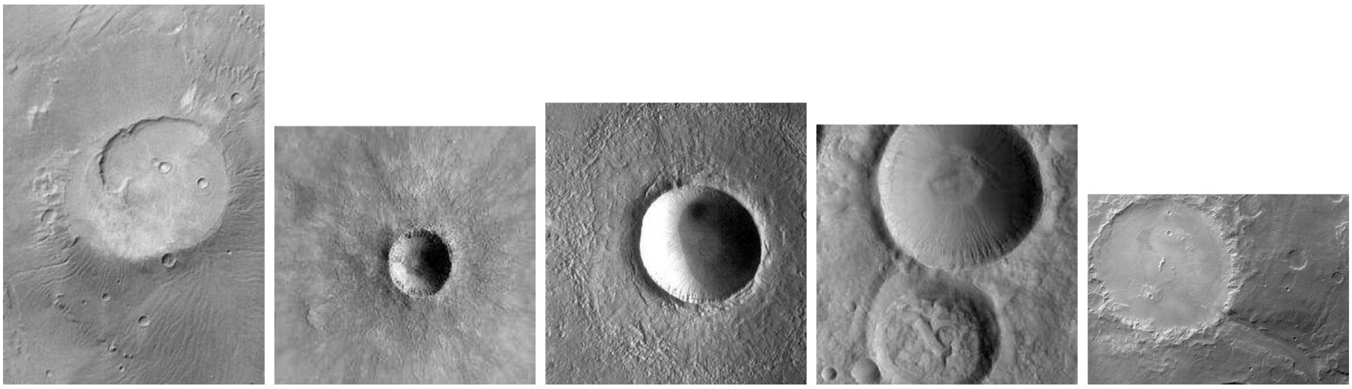
4. Experimentation

In the experimental phase, we have worked with a total of 10 different images (Fig. 10).

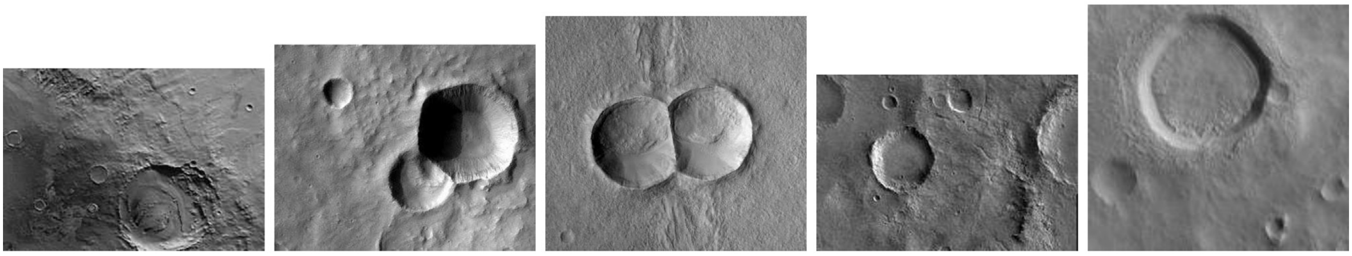
4.1. Descriptions obtained

The relevancy thresholds were fixed to T_{SS} = 0.5 and T_{SP} = 0.8 as an example of application. These parameters can be modified by the expert depending on the level of detail that she/he requires in the final report. We obtained the following linguistic descriptions for each image.

- Image 1 → *"The image contains many medium circles and one big circle. The big circle is in the center of the image"*



(a) Image 1 (284x413) (b) Image 2 (221x219) (c) Image 3 (398x429) (d) Image 4 (225x224) (e) Image 5 (639x465)



(f) Image 6 (278x194) (g) Image 7 (391x308) (h) Image 8 (400x359) (i) Image 9 (273x184) (j) Image 10 (160x151)

Fig. 10. The ten analyzed images of the Mars' surface and their resolution (in pixels).

- Image 2 → *"The image does not contain big circles but it contains one medium circle"*
- Image 3 → *"The image contains one big circle in the center"*
- Image 4 → *"The image contains two big circles and two medium circles. The two big circles are in the bottom center and in the top center parts of the image"*
- Image 5 → *"The image contains one big circle, various small circles and three medium circles. The big circle is in the center left part of the image"*
- Image 6 → *"The image contains various medium circles and one big circle. The big circle is in the bottom center part of the image"*
- Image 7 → *"The image contains one big circle and two medium circles. The big circle is in the center right part of the image"*
- Image 8 → *"The image contains two big circles. The two big circles are in the center of the image"*
- Image 9 → *"The image contains one big circle and three medium circles. The big circle is in the center right part of the image"*
- Image 10 → *"The image contains one big circle and three medium circles. The big circle is in the top center part of the image"*

4.2. Evaluation of the results

Assessing the performance of a system which aims to describe data using NL is a challenging task. The meaning of NL sentences is determined by its context of use, including the personal experience of the writer and the reader [33]. In order to contribute to solve this problem, we have used a straightforward strategy: we have used the computer to create a set of sentences to cover a simple application domain.

In order to asses the matching of the obtained descriptions with the test images, we have done a survey counting with the help of a set of people. The test consisted of showing them five different descriptions generated by the application for each of the ten images. These five different descriptions were chosen among those sequences which got the best validity degrees, including, of course, the best one which is considered the actual output of our system. Since we are trying to model the behavior of a non-expert observer, we asked a total of 22 different people without knowledge about Martian geology to choose the best among these five descriptions.

Table 1
Percentage of agreement of the human observers with our application (*System agreement*) and percentage of agreement of these human observers among them (*People agreement*).

Image	System agreement (%)	People agreement (%)
1	59.1	37.7
2	54.5	41.1
3	72.7	56.2
4	90.9	82.3
5	54.5	33.8
6	59.1	45.9
7	54.5	34.6
8	77.3	63.2
9	18.2	20.8
10	54.5	37.7
All images	59.5	45.3

In Table 1, we show, for each image, the average percentage of agreement of these people with the computational application (*System agreement*), i.e., the percentage of times that the choice of the people is the same as the description obtained by our application. Also, we show the average percentage of agreement of these people among them (*People agreement*), i.e., the average percentage of times that the choice of each person is the same as the description obtained by the rest of the people. The global average values for all the images are also represented at the bottom.

The *system agreement* gets an average for all the images of 59.5% which is indeed greater than the average *people agreement* (45.3%). The highest values correspond to the images number 4, 8, and 3, which have a low number of circles easy to identify. Moreover, the *system agreement* is greater than the *people agreement* for all of the images except the number 9, which means that each person agrees, in average, more with our application than with the descriptions chosen by the rest of the people. This fact can be explained focusing in Fig. 3, where the four recognized circles of the image number 9 are represented. The description made by the application based on these four circles is: *"The image contains one big circle and three medium circles. The big circle is in the center right part of the image"*. However, a human observer who tends to include the big circle in the center right part of the image, tends also to include the circle at the top left corner that is not recognized by the *DAQ Module*. Therefore, not only is the description obtained by our report generator extremely dependent on the results of the *DAQ Module*, but also on the subjectivity of the people.

4.3. Exploring other relevancy thresholds

To proof the effectiveness of relevance and relevancy thresholds, we have modified the parameters T_{SS} and T_{SP}, checking the evolution of the sentences obtained in final reports. Then, the results of this experimental process applied to three of the ten analyzed images are explained as follows:

- Image 2:
  Original: "The image does not contain big circles but it contains one medium circle."
  Increasing T_{SS} from 0.5 to 0.7: "The image does not contain big circles."
  Decreasing T_{SP} from 0.8 to 0.7: "The image does not contain big circles but it contains one medium circle. The medium circle is in the center of the image."
- Image 4:
  Original: "The image contains two big circles and two medium circles. The two big circles are in the bottom center and in the top center parts of the image."
  Increasing T_{SS} from 0.5 to 0.7: "The image contains two big circles. The two big circles are in the bottom center and in the top center parts of the image."
  Decreasing T_{SP} from 0.8 to 0.7: "The image contains two big circles and two medium circles. The two big circles are in the bottom center and in the top center parts of the image. The two medium circles are in the bottom left part of the image."
- Image 9:
  Original: "The image contains one big circle and three medium circles. The big circle is in the center right part of the image."
  Increasing T_{SS} from 0.5 to 0.7: "The image contains one big circle. The big circle is in the top center part of the image."
  Decreasing T_{SP} from 0.8 to 0.7: "The image contains one big circle and three medium circles. The big circle is in the center right part of the image. The three medium circles are in the top left and center left parts of the image."

As we can see in the previous results, when a relevancy threshold is increased, the *Expression Module* omits more information, including only the most relevant information. On the other hand, when a relevancy threshold is decreased, the *Expression Module* includes more information, specifying details less relevant for the expert.

5. Concluding remarks

This paper present our last contributions to the field of generating linguistic descriptions of data:

- We have explored the multidimensionality of the domain of existence of a CP.
- We have developed upon the use of the relevance to choose the most suitable linguistic expressions.
- We have presented a new version of *Report Generator* based on the concept of Fuzzy Tree of Choices.
- We have described a computational application able to create customizable linguistic reports about detected circular structures such as volcanoes or meteorite impacts.

In human generated linguistic reports, the use of NL depends on the application context and specifically of the writer experience and intentions. Here, the chosen linguistic expressions should capture the subjectivity of the human beings participating in the process of designing the computational system, namely, the expert on Martian geology that will provide the functional requirements and the designer that will try to implement that functionality.

Of course, this approach can also be applied to different fields. The difficulties of each application will depend on the complexity of the data and of the complexity of the desired linguistic reports.

In future works in this application, we could improve the *DAQ Module* in order to recognize different geological structures and we will develop new reports to provide more complex descriptions of the Mars' surface.

Acknowledgment

This work has been funded by the Spanish Government (MICINN) under project TIN2011-29827-C02-01.

References

[1] P.R. Christensen, J.L. Bandfield, J.F. Bell, N. Gorelick, V.E. Hamilton, A. Ivanov, B.M. Jakosky, H.H. Kieffer, M.D. Lane, M.C. Malin, T. McConnochie, A.S. McEwen, H.Y. McSween, G.L. Mehall, J.E. Moersch, K.H. Nealson, J.W. Rice, M.I. Richardson, S.W. Ruff, M.D. Smith, T.N. Titus, M.B. Wyatt, Morphology and composition of the surface of Mars: Mars Odyssey THEMIS results, Science 300 (2003) 2056–2061.

[2] D.E. Smith, M.T. Zuber, S.C. Solomon, R.J. Phillips, J.W. Head, J.B. Garvin, W.B. Banerdt, D.O. Muhleman, G.H. Pettengill, G.A. Neumann, F.G. Lemoine, J.B. Abshire, O. Aharonson, C. David, S.A. Brown, A.B. Hauck, P.J. Ivanov, H.J. McGovern, T.C.Duxbury Zwally, The global topography of mars and implications for surface evolution, Science 284 (1999) 1495–1503.

[3] R.C. Gonzalez, R.E. Woods, Digital Image Processing, 2nd ed., Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 2001.

[4] W. Burger, M.J. Burge, Digital Image Processing: An Algorithmic Introduction using Java, Springer, New York, NY (USA), 2007.

[5] A. Eichhorn, D. Girimonte, A. Klose, R. Kruse, Soft computing for automated surface quality analysis of exterior car body panels, Applied Soft Computing 5 (2005) 301–313.

[6] R. Castillo-Ortega, J. Chamorro-Martinez, N. Marin, D. Sanchez, J. Soto-Hidalgo, Describing images via linguistic features and hierarchical segmentation, in: Proceedings of the 2010 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), Barcelona, Spain, 2010, pp. 1104–1111.

[7] S. Jaime-Castillo, J. Medina, D. Sanchez, Using FORDBMS for the linguistic description of images, in: 2010 IEEE International Conference on Fuzzy Systems (FUZZ), 2010, pp. 1–8.

[8] L.A. Zadeh, From computing with numbers to computing with words – from manipulation of measurements to manipulation of perceptions. I: fundamental theory and applications, IEEE Transactions on Circuits and Systems 45 (1) (1999) 105–119.

[9] L.A. Zadeh, A new direction in AI. towards a computational theory of perceptions of measurements to manipulation of perceptions, AI Magazine 22 (2001).

[10] L.A. Zadeh, Toward human level machine intelligence – is it achievable? The need for a paradigm shift, IEEE Computational Intelligence Magazine (2008).

[11] S. Mendez-Nunez, G. Trivino, Combining semantic web technologies and computational theory of perceptions for text generation in financial analysis, in: Proceedings of the 2010 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), Barcelona, Spain, 2010, pp. 953–960.

[12] D. Sanchez-Valdes, L. Eciolaza, G. Trivino, Linguistic description of human activity based on mobile phone's accelerometers, in: Ambient Assisted Living and Home Care, Springer LNCS, 2012, pp. 346–353.

[13] A. Alvarez-Alvarez, G. Trivino, Linguistic description of the human gait quality, Engineering Applications of Artificial Intelligence 26 (1) (2013) 13–23.
[14] L. Eciolaza, G. Trivino, B. Delgado, J. Rojas, M. Sevillano, Fuzzy linguistic reporting in driving simulators, in: Proceedings of the 2011 IEEE Symposium on Computational Intelligence in Vehicles and Transportation Systems, Paris, France, 2010, pp. 30–37.
[15] L. Eciolaza, G. Trivino, Linguistic reporting of driver behavior: Summary and event description, in: Proceedings of the 11th International Conference on Intelligent Systems Design and Applications (ISDA), Córdoba, Spain, 2011, pp. 148–153.
[16] G. Trivino, A. Sanchez, A.S. Montemayor, J.J. Pantrigo, R. Cabido, E.G. Pardo, Linguistic description of traffic in a roundabout, in: Proceedings of the 2010 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), Barcelona, Spain, 2010, pp. 2158–2165.
[17] A. Alvarez-Alvarez, D. Sanchez-Valdes, G. Trivino, Automatic linguistic report about the traffic evolution in roads, Expert Systems with Applications 39 (2012) 11293–11302.
[18] A. Alvarez-Alvarez, D. Sanchez-Valdes, G. Trivino, Automatic linguistic description about relevant features of the Mars' surface, in: Proceedings of the 11th International Conference on Intelligent Systems Design and Applications (ISDA), Córdoba, Spain, 2011, pp. 154–159.
[19] E. Reiter, An architecture for data-to-text systems, in: Proceedings of the Eleventh European Workshop on Natural Language Generation (ENLG' 07), 2007, pp. 97–104.
[20] L.A. Zadeh, The concept of a linguistic variable and its application to approximate reasoning-I, Information Sciences 8 (1975) 199–249.
[21] L.A. Zadeh, The concept of a linguistic variable and its application to approximate reasoning-II, Information Sciences 8 (1975) 301–357.
[22] L.A. Zadeh, The concept of a linguistic variable and its application to approximate reasoning-III, Information Sciences 9 (1975) 43–80.
[23] K. Popper, J. Eccles, The Self and Its Brain, Springer-Verlag, 1977.
[24] M. Sonka, V. Hlavac, R. Boyle, Image Processing, Analysis and Machine Vision, Chapman - Hall Computing, 2006.
[25] S. Etemad, T. White, An ant-inspired algorithm for detection of image edge features, Applied Soft Computing 11 (2011) 4883–4893.
[26] B. Karasulu, S. Korukoglu, A simulated annealing-based optimal threshold determining method in edge-based segmentation of grayscale images, Applied Soft Computing 11 (2011) 2246–2259.
[27] Z. Ji, Y. Xia, Q. Chen, Q. Sun, D. Xia, D. Feng, Fuzzy c-means clustering with weighted image patch for image segmentation, Applied Soft Computing (2012) 1659–1667.
[28] B. Sowmya, B. Rani, Colour image segmentation using fuzzy clustering techniques and competitive neural network, Applied Soft Computing 11 (2011) 3170–3178.
[29] G. Maheswasi, K. Ramar, D. Manimegalai, V. Gomathi, An adaptive region based color texture segmentation using fuzzfied distance metric, Applied Soft Computing 11 (2011) 2916–2924.
[30] J. Illingworth, J. Kittler, A survey of the Hough transform, Computer Vision, Graphics, and Image Processing 44 (1988) 87–116.
[31] E.H. Ruspini, A new approach to clustering, Information and Control 15 (1969) 22–32.
[32] M. Delgado, D. Sánchez, M. Vila, Fuzzy cardinality based evaluation of quantified sentences, International Journal of Approximate Reasoning 23 (2000) 23–66.
[33] M.A.K. Halliday, M.I.M. Matthiessen, Construing Experience through Meaning: A Language-based Approach to Cognition, Continuum; Study ed edition (June 3, 2006), 1999.

Daniel Sanchez-Valdes received his B.S. degree (2008) and his M.S. degree (2011) both in Industrial Engineering from the University of Oviedo (Spain). In 2012, he received the M.S. degree in Soft Computing and Intelligent Data Analysis jointly organized by the European Centre for Soft Computing and the University of Oviedo. He is currently working toward the Ph.D. degree in Computer Science with the same university. Since September 2011 he has been Research Assistant with the "Computing with Perceptions" research unit at the European Centre for Soft Computing. He has published papers in international journals and peer-reviewed conferences, and he has received the Best Special Session Student Paper Award in the 11th International Conference on Intelligent Systems Design and Applications (Córdoba, Spain, 2011). His main research interests are focused on the generation of linguistic descriptions from non-linguistic data, such as sensor signals or database systems.

Alberto Alvarez-Alvarez received his M.S. degree in Physics (2008) from the Universidad Autónoma de Madrid (Spain). In 2010, he received the M.S. degree in Soft Computing and Intelligent Data Analysis jointly organized by the European Centre for Soft Computing and the University of Oviedo. Finally, in 2012, he received the Ph.D. Degree in Computer Science from the same university. Since September 2008, he has been Research Assistant with the "Computing with Perceptions" research unit at the European Centre for Soft Computing. Since January 2013, he is Postdoctoral Researcher in the same research unit. He has published more than fifteen papers in international journals and peer-review conferences. His main research interests include the use of sensors to obtain and manage data from the physical context, analyze them using Soft Computing techniques and produce accurate models close to natural language. Mr. Alvarez Alvarez's three contributions have received the Best Special Session Student Paper Award in the 11th International Conference on Intelligent Systems Design and Applications (Córdoba, Spain, 2011), the IEEE Computational Intelligence Society 2011 Outstanding Paper Award at the 5th IEEE International Workshop on Genetic and Evolutionary Fuzzy Systems (Paris, France, 2011), and the Best Presentation Award at the International Open Workshop Fuzziness and Medicine (Mieres, Spain, 2011).

Gracian Trivino received his M.S. and Ph.D. degrees in Electrical Engineering from the Polytechnic University of Madrid (Spain), in 1982 and 2000, respectively. From 1982 to 1999 he occupied several positions managing R&D projects in industry. His research areas covered during this professional activity include Industrial Control, Nuclear Power Plants Control, Telecommunications, Avionics, and Software Engineering. From 1999 to 2006 he was Associate Professor with the Computer Science School of the Polytechnic University of Madrid. Research topics covered were Intelligent building control, Control of physical processes, Electronic instrumentation, Photonic devices, Development of computational systems, and Architectures for robotics. Since January 2007 he is Principal Researcher with the European Centre for Soft Computing. During these years, he has participated in 20 public funded projects, has written 20 papers in journals, has registered seven patents and has sent more than 50 contributions to peer-reviewed scientific congresses. His current research includes the development of Zadeh's Computational Theory of Perceptions toward linguistic description of complex phenomena.

5.2. Computational Perceptions of uninterpretable data. A case study on the linguistic modeling of human gait as a quasi-periodic phenomenon

D. Sanchez-Valdes, and G. Trivino. “Computational Perceptions of uninterpretable data. A case study on the linguistic modeling of human gait as a quasi-periodic phenomenon”. *Fuzzy Sets and Systems*, Vol. 253, pp. 101–121, 2014.

Available online at www.sciencedirect.com
ScienceDirect
Fuzzy Sets and Systems 253 (2014) 101–121
FUZZY sets and systems
www.elsevier.com/locate/fss

Computational Perceptions of uninterpretable data. A case study on the linguistic modeling of human gait as a quasi-periodic phenomenon

Daniel Sanchez-Valdes *, Gracian Trivino

European Centre for Soft Computing, C/Gonzalo Gutierrez Quirós, s/n, 33600 Mieres, Asturias, Spain

Received 10 June 2013; received in revised form 22 October 2013; accepted 17 December 2013

Available online 24 December 2013

Abstract

As a contribution to the field of the Zadeh's Computational Theory of Perceptions, this paper deals with generating linguistic models of those phenomena that the designer cannot completely recognize at running time. We explore the possibility of generating linguistic descriptions including in their meaning the fact the complete interpretation of input data is not feasible. With this purpose, we extend our previous research in this field by introducing the concept of Computational Perception of uninterpretable data. This extension provides more expressiveness and flexibility to the currently available resources. We define a linguistic variable that explains the degree with which phenomena suit the model, i.e., if input signals can be interpreted. In order to demonstrate the advantages of this contribution, we apply it to create a definition of fuzzy set of quasi-periodic phenomena. We use these new linguistic models to analyze and linguistically describe some characteristics of the human gait, as a case study of quasi-periodic phenomenon.

Keywords: Uninterpretable data; Computing with perceptions; Linguistic modeling; Quasi-periodic signals; Gait analysis

1. Introduction

Human beings use Natural Language (NL) to organize the structure of their experience and to communicate with others [1]. The experience acquired in our everyday life environment using NL allows us to describe our perceptions to others by means of meaningful linguistic expressions that we share with them in specific contexts.

Many times, due to the limits of our personal experience, we do not have available all the resources needed to recognize and describe the relevant details of specific input data. In order to communicate the uncertainty produced by this lack of experience, we use expressions like "as far as I know...", remarking to the listener the lack of a complete knowledge about the meaning of input data. For example, after exploring a damaged car, the expert could tell us: "It seems to be a problem with the battery or the starter engine, but it may be other cause". Obviously, we

* Corresponding author. Tel.: +34 985 45 65 45.
E-mail addresses: daniel.sanchezv@softcomputing.es (D. Sanchez-Valdes), gracian.trivino@softcomputing.es (G. Trivino).

0165-0114/$ – see front matter © 2014 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.fss.2013.12.010

understand that he/she is not very sure about the diagnosis. His/her statement indicates that the expert's experience is not enough to clearly recognize the cause of the car bad functioning.

There is a demand of computational applications able to convert data into relevant linguistic descriptions. According to each situation type, these tools should interpret and represent data in an understandable way, and therefore, generate descriptions so useful as possible to achieve the specific user's goals. Zadeh's principle of incompatibility [2] formulates that the more complex the phenomena, the more difficult to design precise computational models. In order to cope with this drawback, human beings use NL to build linguistic models of perceived phenomena in their environment.

In general, because the inevitable incompleteness of computational models, when monitoring complex phenomena, we only can recognize input data with a certain matching degree. This paper faces with the challenge of enriching the meaning of automatic linguistic descriptions of phenomena in order to contribute to overcome this limitation.

In recent years, there has been a growing interest for modeling approaches based on Fuzzy Logic due to its ability for linguistic concept modeling and system identification. On the one hand, semantic expressiveness, using linguistic variables and rules, is quite close to natural language, which reduces the effort of expert knowledge extraction. On the other hand, being universal approximators [3] fuzzy inference systems are able to perform nonlinear mappings between inputs and outputs. Thanks to these advantages, Fuzzy Logic has been successfully applied in classification [4,5], regression [6,7], control [8,9], system modeling [10,11] and linguistic summarization [12,13], achieving a good interpretability accuracy trade-off [14].

In this context, the idea of *uninterpretable data* arises naturally. We define *uninterpretable data* as input data that do not fully suit with the available model of the monitored phenomenon. The lack of complete knowledge is a typical characteristic of expert systems and data-driven systems, including rule base systems that combine expert rules with induced ones [15,16]. According to [17], "completeness means that for any possible input vector, at least one rule is fired, there is no inference breaking". The lack of complete knowledge or completeness varies depending on the applications. In automatic applications like fuzzy controllers, the rule base should be complete. However, in applications that involve interaction with humans, the interpretability is more appreciated than the completeness. In this case, the higher the number of rules, the lower the system interpretability. In classical fuzzy modeling, two options mainly exist to deal with this problem: to complete the rule base or to use a default rule [18]. Mencar and Fanelli define completeness as "a property of deductive systems that has been used in the context of Artificial Intelligence to indicate that the knowledge representation scheme can represent every entity within the intended domain" [19].

Quasi-periodic phenomena are a type of complex phenomena that provide signals with repetitive temporal patterns but including some variations in period and amplitude. Examples of this type of phenomena are electrocardiograms, accelerations produced during the human gait, vibrations of musical instruments, etc. They are good examples of phenomena that either cannot be modeled, or we do not want to model, with absolute precision. Popular approaches to deal with quasi-periodic signals vary from wavelets transform [20], classical curve fitting methodologies to Hidden Markov Models [21] and Neural Networks [22,23].

Our long term research line is based on the Computational Theory of Perceptions (CTP) introduced by Zadeh [24–26]. In previous works, we have developed computational systems able to generate linguistic descriptions of different types of phenomena. For example, we have generated assessing reports in truck driving simulators [27], reports about traffic evolution in roads [28], about the relevant features of the Mars' surface [29] and linguistic descriptions about visual double stars [30].

When Fuzzy Logic is combined with Finite State Machines, a more fluent modeling process is achieved. Fuzzy Finite State Machines (FFSM) are specially useful tools to model dynamical processes that are time-dependent, becoming an extension of classical Finite State Machines [31,32]. The main advantage of FFSMs is their ability to handle imprecise and uncertain data, which are inherent to real-world phenomena, in the form of fuzzy states and transitions. The theoretical basics of FFSMs were established in [33] and later developed in [34–36]. In previous research, we have learned that FFSMs are suitable tools to model quasi-periodic phenomena. We have applied FFSMs for pattern recognition tasks such as human gait recognition [37,38] and gesture recognition [39,40] based on accelerometer data. We used an FFSM to fuse body posture and WiFi positioning [41].

The context of our contribution can be defined using the following quote from Klir and Yuan [42]: "uncertainty in any problem-solving situation is a result of some information deficiency. Information (pertaining to the model within which the situation is conceptualized) may be incomplete, imprecise, fragmentary, no fully reliable, vague, contradictory, or deficient in some other way. In general, various information deficiencies may result in different

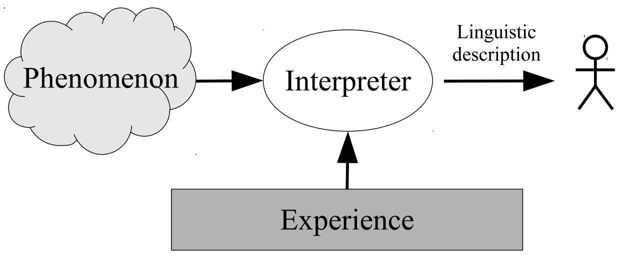


Fig. 1. Basic architecture of a computational system for generating linguistic descriptions of phenomena.

types of uncertainty". Many other approaches have dealt with modeling uncertainty. One of the most complete is the RIMER method [43]. This method deals, on the one hand, with the uncertainty associated with the consequents of rules. Due to the inaccuracy or vagueness existing in information provided by experts, the consequents in a rule-based system can include some uncertainty that this method models with a "belief-degree". Our study does not focus in this type of uncertainty. According with the perspective of RIMER method, here we work with consequents with belief degree = 1. On the other hand, the RIMER method deals with situations in which the input information is modeled by giving different degrees of importance to each rule and to each attribute. Here, we do not use this possibility; we focus our study in linguistically describing the degree to which the model fits to the input data, i.e., describing the lack of the model to cover the whole domain of existence of the monitored phenomenon. The goal is to enrich the linguistic descriptions of phenomena by communicating to the user the ability or inability of the model to recognize the input data. Thus, the main novel contributions of this paper are summarized as follows:

- To improve the potential of meaning of current computational applications for linguistic description of data. Section 2 extends our contribution to the field of Computational Theory of Perceptions by introducing the concept of Computational Perceptions of uninterpretable data.
- To present a definition of fuzzy set of quasi-periodic signals. In Section 3, Computational Perceptions of uninterpretable data are used to design new models of quasi-periodic phenomena. We use FFSMs to produce more versatile and interpretable models of this type of phenomena.
- To contribute to increase the available computational tools for human gait analysis. Section 4 shows how to generate linguistic descriptions of the human gait temporal evolution taking into account the fact that the signal is not always completely interpretable. Some experimental results are discussed in Section 5. In Section 6, we develop upon our previous research in human gait modeling by automatically adapting the model even when the subject performs important changes in the style of walking, regarding both speed and applied force. Section 7 validates this enhanced model.

2. Linguistic description of complex phenomena

Perceptions are "physical sensations interpreted in the light of experience" [44]. Inspired in this definition, Fig. 1 illustrates the general diagram of a computational system able to generate linguistic descriptions of phenomena. The *Interpreter* module gathers input data from a given monitored *Phenomenon* and uses computational models supported by the available *Experience* to recognize meaningful patterns. The output is a *Linguistic description* made up of several linguistic expressions properly combined for describing the meaning of the perceptions to be communicated to specific users in specific application contexts.

2.1. Computational Perception (CP)

A CP is a tuple (A, W) described as follows:

$A =$ $(a_0, a_1, a_2, \ldots, a_n)$ is a vector of linguistic expressions (words or sentences in NL) that represent the whole linguistic domain of CP. The components of A are defined by the designer in accordance with the most suitable sentences from the typically used ones in the application domain of language. Thus, each component a_i is the most suitable linguistic value of CP in each situation of the phenomenon with specific granularity degree. The greater the number of expressions the more granularity and precision can be achieved. Sentences can be either simple, e.g., a_i = *"The temperature is quite high"*, or complex, a_i = *"Today the weather is better than the last days"*. In this paper we extend our previous definition of CP by introducing a_0 as an additional linguistic expression which indicates that the current perception does not match with any of the other available expressions, e.g., *"The available model cannot explain completely the current situation"*. The goal of this additional expression is to model the lack of interpretability of the signal, providing information of those parts of the signal that we cannot explain with the available model. When the whole domain that defines the CP is well-defined, we do not care about uninterpretable data. However, when the definition of the domain reveals certain ignorance because of a lack of complete knowledge, as it usually happens in most real-world applications, handling such uninterpretable data becomes essential. The linguistic expressions a_i will be used to generate linguistic reports about the monitored phenomenon.

$W =$ $(w_0, w_1, w_2, \dots, w_n)$ is a vector of validity degrees $w_i \in [0, 1]$ assigned to each a_i. In the application context, w_i represents the suitability of a_i to describe the current perception. The designer chooses the components of A by trying to cover all the possible values of CP, i.e., the total validity should be distributed among all linguistic labels. Therefore, typically, the components of A are attached to fuzzy sets forming strong fuzzy partitions [45] in the universe of discourse of CP, i.e., $\sum_{i=0}^{n} w_i = 1$. It is worth remarking that, usually, the output of a set of fuzzy rules is obtained after normalization and the information about the non-interpretability of data is lost. The value of w_0 represents the degree up to which the current input value is uninterpretable. It is calculated as $w_0 = 1 - \sum_{i=1}^{n} w_i$.

For example, let us suppose that we want to design a CP about the weather evaluation during the last month. Provided a set of input data characterizing the weather (temperature, humidity, raining rate, etc.) during that period, we could have a set of available linguistic expressions A which describes the weather as *very bad, bad, normal, good* or *very good*. After applying, e.g., a set of fuzzy rules, we may infer that the weather was *very bad* ($w_1 = 0.1$), *bad* ($w_2 = 0.5$), *normal* ($w_3 = 0$), *good* ($w_4 = 0$) and *very good* ($w_5 = 0$). Possible reasons that produce that $\sum_{i=1}^{n} w_i \neq 1$ could be the lack of experience when designing the model, the presence of noise in the signal that difficults its interpretation, errors in data acquisition systems, etc. One way to represent this lack of interpretability in the available data is calculating w_0 as $1 - \sum_{i=1}^{n} w_i$, obtaining that $w_0 = 0.4$. Thus, we could describe the weather during the last month with the following sentence: *"As far as I know, the weather during the last month was bad"*, where the expression *"As far as I know"* denotes a lack of confident into the meaning of the sentence when describing the weather.

2.2. Perception Mapping (PM)

We use PMs to combine or aggregate CPs. A PM is a tuple (U, y, g, T) where:

U is a vector of input CPs, $U = (u_1, u_2, \dots, u_m)$, where u_i are tuples (A_i, W_i) and m the number of input CPs.

y is the output CP, $y = (A_y, W_y)$. That is to say PM: $u_1 \times u_2 \times \dots \times u_m \to y$.

g is an aggregation function $W_y = g(W_1, W_2, \dots, W_m)$, where W_i are the vectors of validity degrees of the m input CPs.
In Fuzzy Logic, many different types of aggregation functions have been developed. For example, g can be implemented using a set of fuzzy rules. As mentioned in the Introduction, the concept of *default rule* has been already studied in the field of sets of fuzzy rules. Here, we use this concept to give value to w_0 in W_y.

T is a text generation algorithm that allows generating the sentences in A_y. In our current approach, T is typically a basic linguistic template, e.g., *"Likely the temperature sensor is wrong"* and/or *"The temperature of this room is* {*high* | *medium* | *low*}*"*.

We use a special type of perception mapping to process the input data of the whole computational model. We call them *first order perception mappings* (1PMs). In this special case the inputs are numerical data obtained either by sensors or obtained from results of a database query. For the sake of homogeneity, we consider these numerical inputs as a very special case of CP where A_y is the set of linguistic expressions of these numbers, e.g., a_i = "25.2", and W is a vector where $w_i = 1$ and $w_j = 0, \forall j \neq i$.

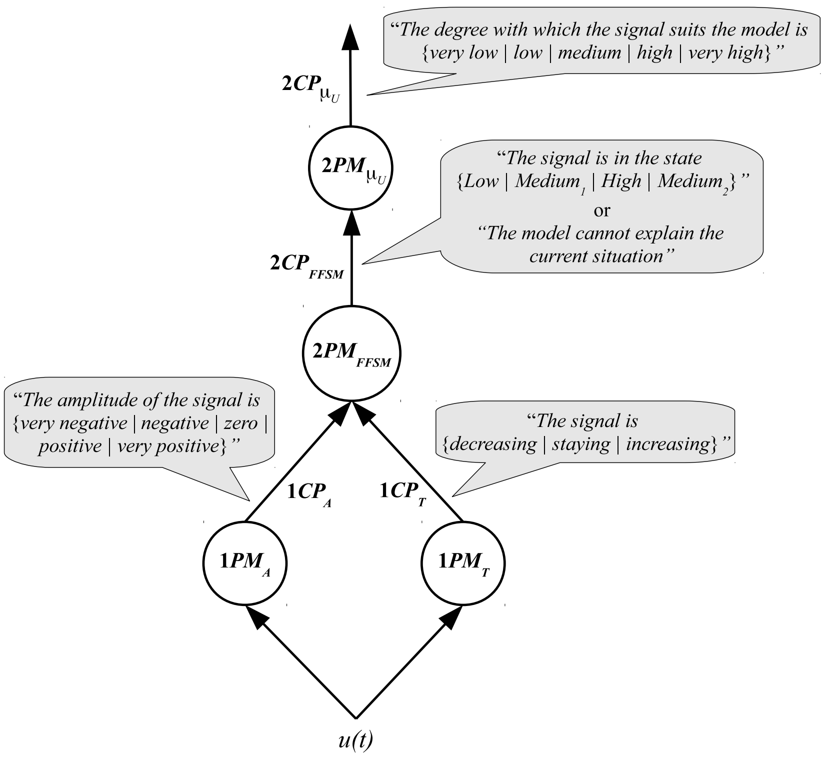


Fig. 2. GLMP for the linguistic description of the $\mu_U(u)$ of a sinusoidal signal.

2.3. *Granular Linguistic Model of Phenomena (GLMP)*

GLMPs consist of networks of PMs (see in Fig. 2 a simple example of GLMP that we explain in next section). We say that output CPs are explained by the PM using a set of input CPs. In the network, each CP covers a specific aspect of the phenomenon with certain granularity degree.

We call *first order computational perceptions* (1CPs) to those obtained from the system input and we call *second order computational perception* (2CPs) to those explained by previous CPs. This classification is inspired on the definition of the three worlds by Popper, namely, the world-1 of *physical objects* (phenomena), the world-2 of the *perceived objects* (1CP) and the world-3 of the *mental objects* built by using the objects in the world-2 (2CP) [46].

By means of using different aggregation functions and different linguistic expressions, the GLMP paradigm allows the designer to model computationally his/her perceptions. Note that, after being instantiated with a set of input data, the GLMP provides a structure that, in medium size applications, could include hundreds of valid sentences.

2.4. *Fuzzy Finite State Machine (FFSM)*

In system identification, designers decide which paradigm they use to represent system models. The *state space* representation is one of the most expressive model structures, where the designer must find out the necessary and sufficient subset of *state variables* to represent the entire state of the system at every time instant [47]. In every case, the designer uses his/her creativity and personal experience to choose the most adequate set of *state variables* depending on the system goals. This set of variables emphasizes the relevant aspects of the system and hides the irrelevant ones.

When the system evolves in time, the current state follows a trajectory in the *state space*. In this context, an FFSM can be considered a particular case of *state space* model that is especially useful when modeling phenomena that evolves in time through a limited number of fuzzy states. This paradigm is used in system modeling to allow experts to build comprehensible linguistic fuzzy models in an easier way. An FFSM is a tuple $\{Q, U, f, Y, g\}$, where:

- Q is the state vector of the system: $(q_0, q_1, q_2, \ldots, q_n)$, with n being the number of states, and q_0 is the initial state of the system.

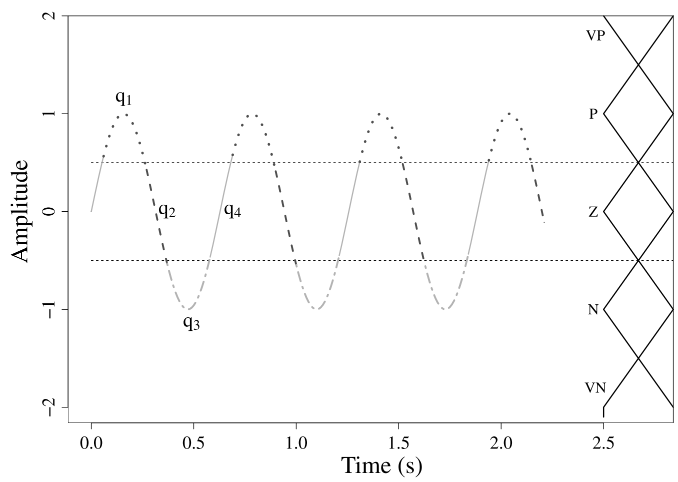


Fig. 3. Sinusoidal signal with related fuzzy labels.

- U is the input vector of the system: $(u_1, u_2, \ldots, u_{n_u})$, with n_u being the number of input variables.
- f is the transition function that calculates the state vector of the system.
- Y is the output vector: $(y_1, y_2, \ldots, y_{n_y})$, with n_y being the number of output variables.
- g is the output function that calculates the output vector of the system.

For a more detailed description of our approach, we refer the interested reader to some application examples on how to implement FFSMs [37,50].

3. Fuzzy sets of quasi-periodic signals

According to the definition of fuzzy set [48], we define a fuzzy set U of quasi-periodic signals by using a membership function $\mu_U(u)$ that provides the degree with which a signal u belongs to U. We introduced this concept in [49] and here we develop upon it by using Computational Perceptions of uninterpretable data.

In order to illustrate the method for creating this type of membership functions $\mu_U(u)$, we use a simple example built around a periodical signal $u(t) = \sin(\omega t)$ that we will modify in amplitude and frequency. Fig. 2 shows a GLMP designed to describe this signal. $1CP_A$ and $1CP_T$ describe the amplitude and trend of the input signal respectively. $2CP_{FFSM}$ shows the state of the signal at each time instant and $2CP_{\mu_U}$ provides the value of $\mu_U(u)$ evaluated along certain temporal window. In the following subsections, we thoroughly describe each PM in this GLMP.

3.1. Signal amplitude ($1PM_A$)

U are the numerical values ($u(t) \in \mathbb{R}$) of the input signal $u(t) = \sin(\omega t)$, where $\omega = 10$ and t takes values in the time interval $t \in [0, 8\pi]$. The period of the signal is, therefore, $K = 2\pi/10$.

y is the output $1CP_A$, which describes the possible values of the signal amplitude with the vector $A = (Very\ negative$ (VN), $Negative$ (N), $Zero$ (Z), $Positive$ (P), $Very\ positive$ $(VP))$.

g is the function that calculates the validity degrees of the output CP. This first example models a quasi-periodic sinusoidal signal that is very simple and symmetrical. Therefore, the linguistic labels used to design the computational perceptions are clearly defined using simple strong fuzzy partitions. The validity degrees of the output CP are obtained by means of five uniformly distributed triangular membership functions (MFs) associated to each NL expression. They are defined by their vertices as follows: $\{VN\ (-\infty, -2, -1), N\ (-2, -1, 0), Z\ (-1, 0, 1), P\ (0, 1, 2), VP\ (1, 2, \infty)\}$. These linguistic labels are represented in Fig. 3, together with the input signal $u(t)$.

T is the text generator that produces linguistic expressions as follows: *"The amplitude of the signal is* {*very negative* | *negative* | *zero* | *positive* | *very positive*}*"*.

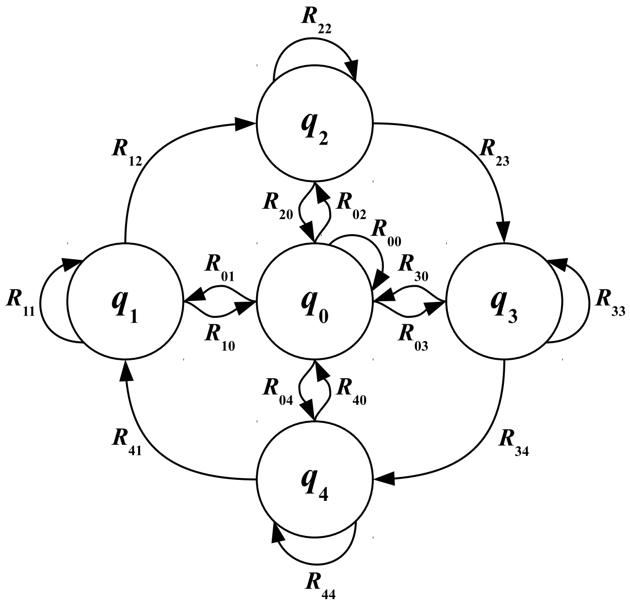


Fig. 4. State diagram of the FFSM for the sinusoidal signal.

3.2. *Signal trend* ($1PM_T$)

U are, again, the numerical values ($u(t) \in \mathbb{R}$) of the input signal.

y is the output $1CP_T$, which describes the possible values of the signal trend with the vector $A =$ (*Decreasing* (*D*), *Staying* (*S*), *Increasing* (*I*)).

g is the output function that calculates the signal trend as $u(t) - u(t-1)$. It is defined with triangular MFs as follows: $\{D\,(-\infty, -0.01, 0), S\,(-0.01, 0, 0.01), I\,(0, 0.01, \infty)\}$.

T produces linguistic expressions as follows: *"The signal is* {*decreasing* | *staying* | *increasing*}*"*.

3.3. *Signal state* ($2PM_{FFSM}$)

In order to describe the temporal evolution of the signal, we have split it into four states (Fig. 3). This $2PM_{FFSM}$ receives the information about the amplitude and trend of the signal at each time instant t and yields the state in $t+1$. It has the following elements:

U are the input CPs: $\{1CP_A, 1CP_T\}$.

y is the output $2CP_{FFSM}$, which describes the possible states of the signal as the states of an FFSM. $A =$ (*Uninterpretable* (q_0), *Low* (q_1), *Medium*$_1$ (q_2), *High* (q_3), *Medium*$_2$ (q_4)), where the "uninterpretable" state (q_0) represents the non-interpretability of the signal.

g is the aggregation function implemented using a set of fuzzy rules that are organized in the form of an FFSM. We have used twelve fuzzy rules, namely, R_{ii} to remain in the state i and R_{ij} to change from state i to state j. Fig. 4 shows, in an illustrative example, how we can use the FFSM to define the possible transitions to change or remain in a state q_i. These fuzzy rules have the following structure:

$$\begin{aligned} R_{ii}:\quad & \text{IF } \big(\textit{State}[t] \text{ is } q_i\big) \text{ AND } (\textit{Input variables constraints})_{ii} \text{ AND } (\textit{Temporal constraints})_{ii} \\ & \text{THEN } \big(\textit{State}[t+1] \text{ is } q_i\big) \\ R_{ij}:\quad & \text{IF } \big(\textit{State}[t] \text{ is } q_i\big) \text{ AND } (\textit{Input variables constraints})_{ij} \text{ AND } (\textit{Temporal constraints})_{ij} \\ & \text{THEN } \big(\textit{State}[t+1] \text{ is } q_j\big) \end{aligned} \tag{1}$$

The antecedent of each rule is composed by the current state of the signal in t, the input variables constraints, and the temporal constraints, which are described below in Section 3.3.1 and Section 3.3.2, respectively. Moreover, in Section 3.3.3, we detail the set of fuzzy rules implemented to model the sinusoidal signal.

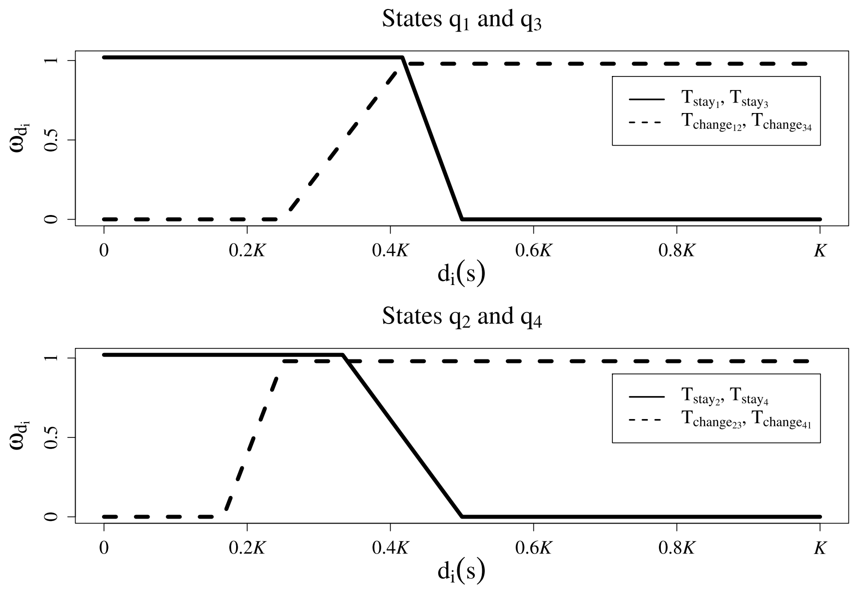


Fig. 5. Times to stay and times to change of states.

T produces linguistic expressions that are adapted depending on the situation. When the signal is in the states q_1, q_2, q_3 or q_4 the template is the following: *"The signal is in the state* {*Low* | *Medium*$_1$ | *High* | *Medium*$_2$}*"*. However, when the signal is in the state q_0, that is, the signal is uninterpretable, the system reports the following expression: *"The model cannot explain the current situation"*. If this occurs, we could analyze the rules R_{i0} that activates the state q_0 to interpret, in a more detailed way, the reason why the model does not fit the signal.

3.3.1. Input variables constraints

Input variables constraints are formed by the conditions of signal amplitude ($2CP_A$) and signal trend ($2CP_T$), e.g., (*Input variables constraints*)$_{22}$ = (*amplitude* is Z) AND (*trend* is D), that represent the input variables constraints to remain in state q_2.

3.3.2. Temporal constraints

Temporal conditions apply constraints over the states duration. We compute numerically the duration d_i of each state q_i as the time that $w_{q_i} > 0$. When the state q_i takes value zero, its duration d_i takes value zero. However, when q_i takes values bigger than zero, its duration d_i increases its value according to the time passed since the last instant of time. The numerical value of d_i allows to calculate the linguistic variable *duration*$_i$, that will be used to apply temporal fuzzy constraints. Thus, we can measure the duration of the phenomenon in each state. We obtain a vector of durations (d_1, d_2, d_3, d_4), and define two temporal conditions:

Time to stay (T_{stay_i}) is the maximum duration that the signal is allowed to stay in state i.

Time to change ($T_{change_{ij}}$) is the minimum duration that the signal must be in state i before changing to state j. This condition helps to filter out very short stays in states when they could be caused by noise in the input signal.

Fig. 5 shows the linguistic labels used to define the temporal constraints. Here, in agreement with our knowledge about the modeled signal, we assign to each state a duration according to its percentage of the period K. In this application, states q_1 and q_3 have the same duration and the linguistic labels can be represented with trapezoidal MFs defined by their vertices as follows: T_{stay_1}, T_{stay_3} $(-\infty, -\infty, \frac{5\cdot K}{12}, \frac{K}{2})$ and $T_{change_{12}}, T_{change_{34}}$ $(\frac{K}{4}, \frac{5\cdot K}{12}, \infty, \infty)$.

On the other hand, states q_2 and q_4 share the same duration and the linguistic labels can be defined by their vertices as follows: T_{stay_2}, T_{stay_4} $(-\infty, -\infty, \frac{K}{3}, \frac{K}{2})$ and $T_{change_{23}}, T_{change_{41}}$ $(\frac{K}{6}, \frac{K}{4}, \infty, \infty)$.

3.3.3. Set of fuzzy rules

According to the structure expressed in (1), the whole rule base that we have designed to model the sinusoidal signal is listed as follows:

R_{11}: IF ($State[t]$ is q_1) AND ($amplitude$ is P) AND ($duration_1$ is T_{stay_1}) THEN ($State[t+1]$ is q_1)
R_{22}: IF ($State[t]$ is q_2) AND ($amplitude$ is Z) AND ($trend$ is D) AND ($duration_2$ is T_{stay_2}) THEN ($State[t+1]$ is q_2)
R_{33}: IF ($State[t]$ is q_3) AND ($amplitude$ is N) AND ($duration_3$ is T_{stay_3}) THEN ($State[t+1]$ is q_3)
R_{44}: IF ($State[t]$ is q_4) AND ($amplitude$ is Z) AND ($trend$ is I) AND ($duration_4$ is T_{stay_4}) THEN ($State[t+1]$ is q_4)
R_{12}: IF ($State[t]$ is q_1) AND ($amplitude$ is Z) AND ($trend$ is D) AND ($duration_1$ is $T_{change_{12}}$) THEN ($State[t+1]$ is q_2)
R_{23}: IF ($State[t]$ is q_2) AND ($amplitude$ is N) AND ($trend$ is D) AND ($duration_2$ is $T_{change_{23}}$) THEN ($State[t+1]$ is q_3)
R_{34}: IF ($State[t]$ is q_3) AND ($amplitude$ is Z) AND ($trend$ is I) AND ($duration_3$ is $T_{change_{34}}$) THEN ($State[t+1]$ is q_4)
R_{41}: IF ($State[t]$ is q_4) AND ($amplitude$ is P) AND ($trend$ is I) AND ($duration_4$ is $T_{change_{41}}$) THEN ($State[t+1]$ is q_1)
R_{01}: IF ($State[t]$ is q_0) AND ($amplitude$ is P) AND ($trend$ is I) THEN ($State[t+1]$ is q_1)
R_{02}: IF ($State[t]$ is q_0) AND ($amplitude$ is Z) AND ($trend$ is D) THEN ($State[t+1]$ is q_2)
R_{03}: IF ($State[t]$ is q_0) AND ($amplitude$ is N) AND ($trend$ is D) THEN ($State[t+1]$ is q_3)
R_{04}: IF ($State[t]$ is q_0) AND ($amplitude$ is Z) AND ($trend$ is I) THEN ($State[t+1]$ is q_4)
R_{i0}: ELSE ($State[t+1]$ is q_0)

The values of $State[t+1]$ are calculated as a weighted average of the individual rules, where the weight of each rule R_{ij} corresponds to its firing degree τ_{ij}. This firing degree is calculated using the minimum for the AND operator and the bounded sum of Łukasiewicz [52] for the OR operator. The activation degree of q_0, i.e., the non-interpretability degree of the input signal, is obtained by means of Eq. (2):

$$w_{q_0}[t+1] = \begin{cases} 1 - \sum_{i=0}^{n} \sum_{j=0}^{n} \tau_{ij} & \text{if } \sum_{i=0}^{n} \sum_{j=0}^{n} \tau_{ij} \leqslant 1 \\ 0 & \text{if } 1 < \sum_{i=0}^{n} \sum_{j=0}^{n} \tau_{ij} \end{cases} \tag{2}$$

The activation degree of the rest of states is calculated by means of Eq. (3):

$$w_{q_j}[t+1] = \begin{cases} \sum_{i=0}^{n} \tau_{ij} & \text{if } \sum_{i=0}^{n} \sum_{j=0}^{n} \tau_{ij} \leqslant 1 \\ \frac{\sum_{i=0}^{n} \tau_{ij}}{\sum_{i=0}^{n} \sum_{j=0}^{n} \tau_{ij}} & \text{if } 1 < \sum_{i=0}^{n} \sum_{j=0}^{n} \tau_{ij} \end{cases} \tag{3}$$

3.4. Matching degree of the signal ($2PM_{\mu_U}$)

This CP deals with describing the signal behavior along a period of time. It calculates the degree with which the input signal suits the model. Here, in addition to obtain a numerical value for the matching degree μ_U, $2CP_{\mu_U}$ provides a linguistic expression of this value. It has the following elements:

- U is the input $2CP_{FFSM}$. The introduction of q_0 in the definition of our FFSM has the goal to represent the impossibility of the model to interpret the input data when the total firing degree of the rules is less than one. This is to say, a poor firing degree of the rules, i.e., the sum of firing degrees is less than one, corresponds to a lack of the model to recognize the current state of the monitored phenomenon.
- y is the output $2CP_{\mu_U}$, which describes the matching degree of the input signal with respect to the modeled one. $A = (Very\ low\ (VL), Low\ (L), Medium\ (M), High\ (H), Very\ high\ (VH))$.
- g is the aggregation function which calculates the matching degree (μ_U) of the signal for a temporal window of n samples. With the "uninterpretable" state (q_0), we can define the μ_U of the output based on its activation degrees ($w_{q_0}[t]$) along a period of time of n samples. Such μ_U is defined as:

$$\mu_U = 1 - \frac{\sum_{t=1}^{n} w_{q_0}[t]}{n} \tag{4}$$

  Thus, μ_U takes values in the range $[0, 1]$ being equal to one if the "uninterpretable" state has not been activated, zero when the "uninterpretable" state has been fully activated, and a real value between zero and one otherwise [41]. The linguistic labels associated to each expression are represented by trapezoidal MFs defined by their vertices as follows: $\{VL\ (-\infty, -\infty, 0.45, 0.5)$, $L\ (0.45, 0.5, 0.6, 0.65)$, $M\ (0.6, 0.65, 0.75, 0.8)$, $H\ (0.75, 0.8, 0.9, 0.95)$, $VH\ (0.9, 0.95, \infty, \infty)\}$.
- T produces linguistic expressions as follows: *"The degree with which the signal suits the model is {very low | low | medium | high | very high}"*.

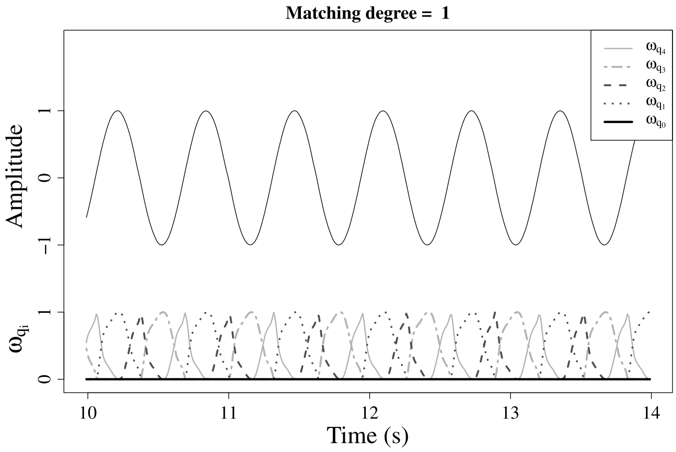


Fig. 6. States and μ_U when the signal is exactly the expected one.

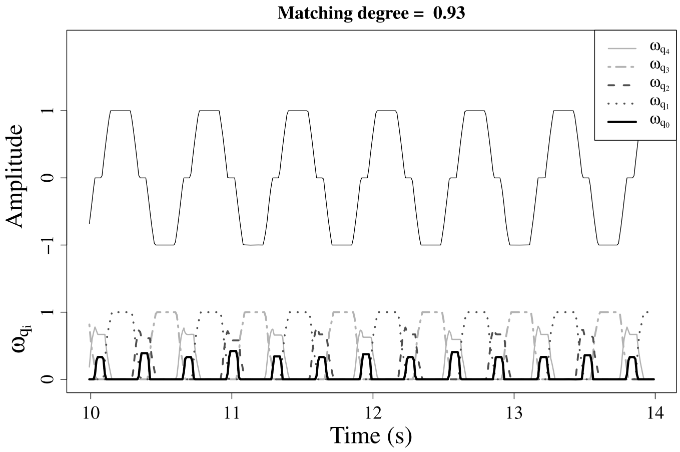


Fig. 7. States and μ_U when the signal is slightly different from the expected one.

3.5. Examples

In order to illustrate how the model works, Figs. 6, 7, 8 and 9 show different behaviors of the previously defined FFSM, depending on the input signal. The GLMP was designed to recognize the states of an input signal of the form $u(t) = \sin(10t)$. In these examples, we want to see how the GLMP works when the input signal varies with respect to the original one. The input signal is represented in the top of each figure and the evolution of each state q_i (w_{q_i}), remarking w_{q_0} with a darker line, is illustrated in the bottom. In addition, we provide the matching degree of the input signal with the model during a temporal window for each example.

3.5.1. Example 1

Fig. 6 shows the fuzzy states evolution when the input signal is exactly the expected one (Eq. (5)), i.e., the set of rules represents perfectly the signal evolution. Therefore, w_{q_0} is equal to zero along the whole lapse of time and μ_U is maximum ($\mu_U(u_1) = 1$).

$$u_1(t) = \sin(10t) \tag{5}$$

3.5.2. Example 2

When the monitored signal is different from the expected one (Eq. (6)), a perturbation or an unforeseen situation uncovered by the set of rules, the μ_U decreases ($\mu_U(u_2) = 0.93$), indicating that the input signal does not correspond to the typical sinusoidal signal that was modeled. Fig. 7 shows an example of this type of situation. We model a trapezoidal wave, similar to the original signal in amplitude and period.

$$u_2(t) = \begin{cases} -1 & \text{if } u(t) \leqslant -0.75 \\ -1 + (u(t) + 0.75) \cdot 2 & \text{if } -0.75 < u(t) \leqslant -0.25 \\ 0 & \text{if } -0.25 < u(t) \leqslant 0.25 \\ 1 + (u(t) - 0.75) \cdot 2 & \text{if } 0.25 < u(t) \leqslant 0.75 \\ 1 & \text{if } u(t) > 0.75 \end{cases} \tag{6}$$

We can see how μ_U indicates that, despite being a very similar signal, the trapezoidal signal does not correspond exactly with the sinusoidal signal. w_{q_0} takes higher values especially in the intermediate zones of the signal, when the trapezoidal signal is zero for short periods of time.

3.5.3. Example 3

Fig. 8 shows the fuzzy states evolution when the amplitude varies along the time (Eq. (7)). When the signal differs from the expected one, in amplitude or period, w_{q_0} takes high values. Otherwise, when the signal behavior is similar to the expected one, w_{q_0} decreases to take low values. In this example, w_{q_0} starts with a high level because the signal

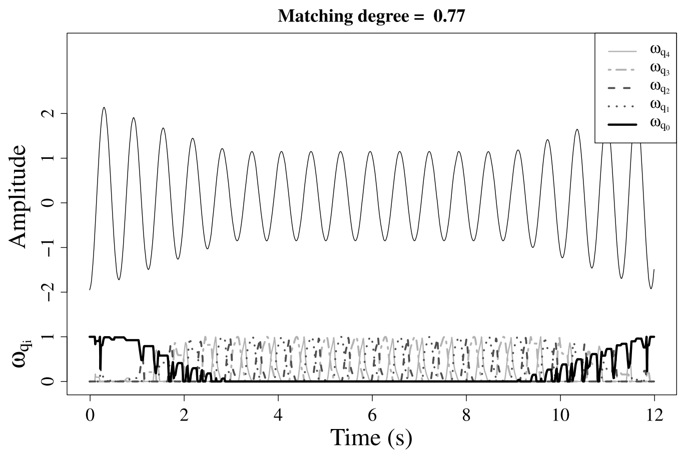


Fig. 8. States and μ_U when the signal varies its amplitude along the time.

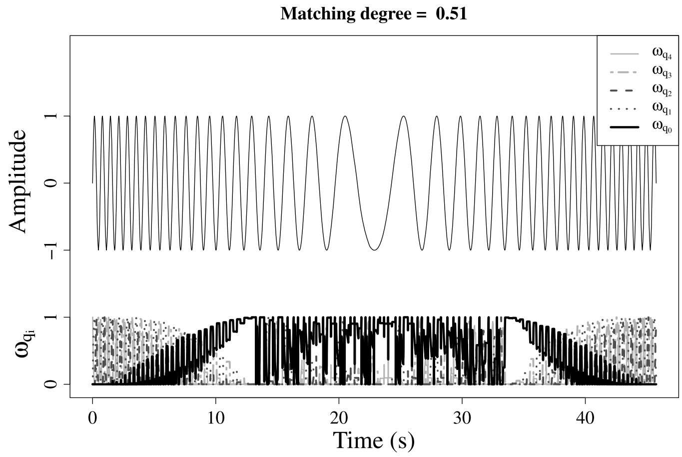


Fig. 9. States and μ_U when the signal varies its period along the time.

amplitude is different from the expected one. When the signal recovers its natural amplitude, w_{q0} decreases to lower values.

In this example $t \in [0, 12]$ and the matching degree of the signal is $\mu_U(u_3) = 0.77$.

$$u_3(t) = \begin{cases} (\frac{-4t}{3} + 5) \cdot \sin(10t) & \text{if } t \leqslant 3 \\ \sin(10t) & \text{if } 3 < t \leqslant 9 \\ (\frac{4(t-9)}{3} + 1) \cdot \sin(10t) & \text{if } t > 9 \end{cases} \tag{7}$$

3.5.4. Example 4

Fig. 9 shows the fuzzy states evolution when the period varies along the time (Eq. (8)). In this case w_{q0} initially takes value zero, but when the signal changes its natural period, w_{q0} increases to higher values. Again, when the period of the signal is similar to the expected one, w_{q0} decreases to take value zero. As we can see in the illustration, even if the signal period changes radically, w_{q0} does not take value one over the time. This is because the input signal continues being a sinusoidal wave and it shares characteristics with the original one. This allows the system to know that the modeled signal is not the expected one but it shares a set of similarities with it, which is different from having a totally different signal that may demand any special action by the controlling system.

In this example, $t \in [0, 45]$ and the matching degree of the signal is $\mu_U(u_4) = 0.51$.

$$u_4(t) = \begin{cases} \sin(10t) & \text{if } t \leqslant 3 \\ \sin((\frac{-5(t-3)}{19.5} + 10) \cdot t) & \text{if } 3 < t \leqslant 22.5 \\ \sin((\frac{5(t-22.5)}{19.5} + 5) \cdot t) & \text{if } 22.5 < u(t) \leqslant 42 \\ \sin(10t) & \text{if } t > 42 \end{cases} \tag{8}$$

4. Case study to human gait modeling

Human gait modeling consists of studying the biomechanics of this human movement aimed at quantifying factors that govern the functionality of the lower extremities. Gait is a complex integrated task that requires precise coordination of the neural and musculoskeletal system to ensure correct skeletal dynamics [53]. Therefore, its analysis can help in the diagnosis and treatment of walking and movement disorders, identification of balance factors, and assessment of clinical gait interventions and rehabilitation programs [54,55].

The gait cycle is a quasi-periodic phenomenon which is defined as the interval between two successive events (usually heel contact) of the same foot [56]. As it can be appreciated in Fig. 10, human gait is characterized by

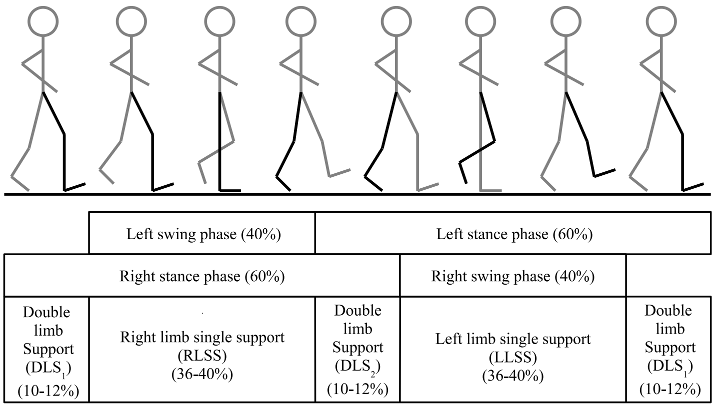


Fig. 10. States and events of a gait cycle.

a stance phase (60% of the total gait cycle), where at least one foot is in contact with the ground, and a swing phase (40% of the total gait cycle), during which one limb swings up to the next heel contact. These phases can be quite different among individuals but when normalized to a percentage of the gait cycle they maintain close similarity, indicating the absence of disorders [57].

In previous works in this research line, we have built a model that linguistically describes the human gait with different levels of granularity [38]. In this work, we showed that, by combining GLMP with FFSM, the designer can create hierarchical models of complex phenomena that work with several input variables while maintaining the model interpretability.

In the currently available model of the human gait, both the definition of rules and the definition of the variables used in the antecedents are the result of an intense process of experimentation done during the last years. A wide range of possibilities have been analyzed to find those variables that reflect the best features that differentiate one state from another while maintaining the model as interpretable as possible.

In this paper, as an extension of this research, we focus on presenting an FFSM that models the Computational Perception of uninterpretable accelerations when we try to recognize the human gait. It is worth to remark that, the goal of this case study is not to present a new model of human gait but using the existing one for applying the new concept. For example, if a person walks very different from the natural way, it is expected that the produced accelerations will not completely suit with the available gait model.

The non-interpretability can be derived from variations in walking speed, force of steps, different footwear and surfaces, etc. In the same way, other sources of uninterpretable data could be errors in measuring data or improper sensor positioning. Our approach consists of identifying the relevant states of the gait based on the accelerations that are produced during the process. We have measured the accelerations using a smartphone placed in the waist and centered in the front of the person. This measure provides the superior–inferior acceleration a_x, the medio-lateral acceleration a_y, and the antero-posterior acceleration a_z at each instant of time. The acquisition frequency is 100 Hz.

We use a linguistic model to represent the subjective designer's perceptions of the human gait process. This model is easily understood and does not require high computational cost.

As in Section 3, here we have designed a simple GLMP (Fig. 11) that describes the human gait states, including the "uninterpretable" state when the model does not recognize properly the current situation. The following subsections detail each PM included in the GLMP.

4.1. Accelerations module amplitude ($1PM_\rho$)

- U are the numerical values ($z_1 \in \mathbb{R}$) of the accelerations module minus the gravity. The accelerations module is calculated as $\rho = \sqrt{a_x^2 + a_y^2 + a_z^2}$.
- y is the output $1CP_\rho$, which describes the possible values of the acceleration module with the vector $A = (\mathit{Small}\,(S), \mathit{Medium}\,(M), \mathit{Big}\,(B))$.

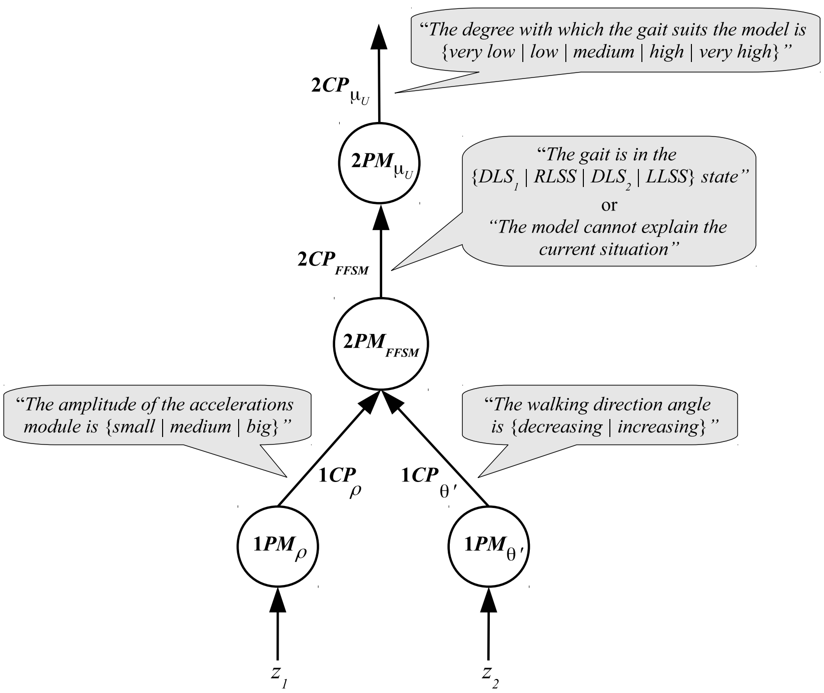


Fig. 11. GLMP for the linguistic description of the human gait phases.

g is the output function that calculates the validity degrees of the output CP. In this second example, we have also used uniformly distributed linguistic labels. In rule-based systems that are linguistically built from expert knowledge it is common to use different paradigms to define the most suitable set of linguistic labels, e.g., one of the available strategies is the use of comparative linguistic expressions [51]. Here, for the sake of simplicity, the validity degrees of the output CP are obtained by means of uniformly distributed trapezoidal MFs forming a strong fuzzy partition. Although we could use a more complex model, thanks to the wide experimentation performed in previous works [38], we found that the use of uniformly linguistic labels is enough to get good results. The linguistic labels associated to each NL expression are defined by their vertices as follows: $\{S\ (-\infty, -\infty, -2, -0.5), M\ (-2, -0.5, 0, 1.5), B\ (0, 1.5, \infty, \infty)\}$.

T produces linguistic expressions as follows: *"The amplitude of the accelerations module is {small | medium | big}"*.

4.2. Trend of the walking direction angle ($1PM_{\theta'}$)

U are the numerical values ($z_2 \in \mathbb{R}$) of the variation of the walking direction angle θ. Notice that θ is the angle between the accelerations module (ρ) and the walking direction.

y is the output $1CP_{\theta'}$, which describes the possible values of the trend of θ with the vector $A = $ (*Decreasing* (*D*), *Increasing* (*I*)).

g is the output function similar to the corresponding to $1PM_\rho$. Here, the two linguistic labels associated to each NL expression are represented with trapezoidal MFs defined by their vertices as follows: $\{D\ (-\infty, -\infty, -0.001, 0.001), I\ (-0.001, 0.001, \infty, \infty)\}$.

T produces linguistic expressions as follows: *"The walking direction angle is {decreasing | increasing}"*.

4.3. Gait states ($2PM_{FFSM}$)

U are the input CPs: $\{1CP_\rho, 1CP_{\theta'}\}$.

y is the output $2CP_{FFSM}$, which describes the possible states of the gait as the states of an FFSM. $A = $ (*Uninterpretable* (q_0), *Double limb support*$_1$ (q_1), *Right limb single support* (q_2), *Double limb support*$_2$ (q_3), *Left limb*

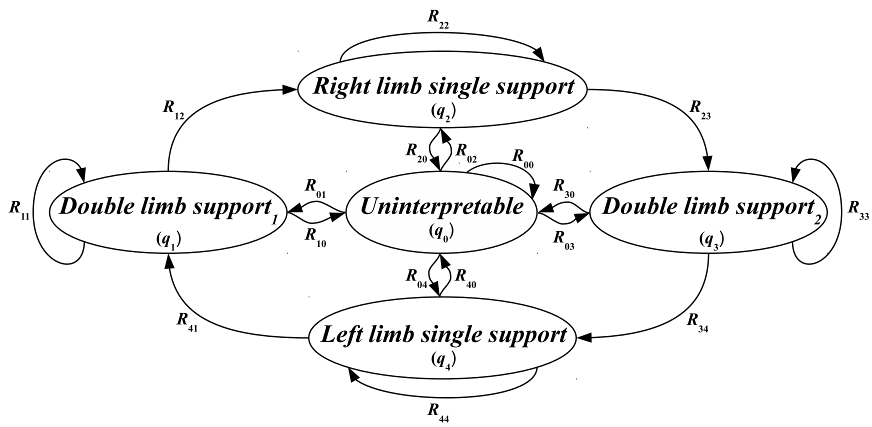


Fig. 12. State diagram of the FFSM for the human gait modeling.

single support (q_4)). Fig. 12 represents the states diagram that shows the transitions allowed in this human gait modeling.

g is the aggregation function which calculates, at each time instant, the next value of the validity degrees for each sentence based on the previous validity degrees and current input CPs. We use an FFSM to implement the $2PM_{FFSM}$, allowing to build a comprehensible linguistic fuzzy model in an easy way.

Temporal constraints T_{stay_i} and $T_{change_{ij}}$ limit the gait to stay in state q_i or to change from state q_i to state q_j, according to the duration d_i. They were defined in accordance with relevant existing works about human gait. Due to the similarities between the states q_1 and q_3, their temporal constraints are the same and they are represented by trapezoidal MFs defined by their vertices as follows: T_{stay_1}, T_{stay_3} $(-\infty, -\infty, 0.2 \cdot K, 0.3 \cdot K)$ and $T_{change_{12}}, T_{change_{34}}$ $(0.05 \cdot K, 0.15 \cdot K, \infty, \infty)$, where K is the period of the gait that initially takes value 1 (1 gait cycle per second). This value is obtained statistically as the average speed at which users walk when they are asked to walk at a regular comfortable speed. In the same way, the states q_2 and q_4 share the temporal constraints which are defined as: T_{stay_2}, T_{stay_4} $(-\infty, -\infty, 0.35 \cdot K, 0.45 \cdot K)$ and $T_{change_{23}}, T_{change_{41}}$ $(0.2 \cdot K, 0.3 \cdot K, \infty, \infty)$.

From the state diagram shown in Fig. 4, we can recognize rules to remain in each state i and rules to change from state i to state j. This set of rules has been designed by extracting the experts' knowledge in human gait and its parameters. Both the variables used in the definition of the antecedents as the number of rules that model the phenomenon is the result of an intense collaboration process with experts. Therefore, the rule base has the following structure:

R_{11}: IF ($S[t]$ is q_1) AND (ρ is B) AND ($duration_1$ is T_{stay_1}) THEN ($S[t+1]$ is q_1)
R_{22}: IF ($S[t]$ is q_2) AND (((ρ is S) OR (ρ is M)) OR (θ' is D)) AND ($duration_2$ is T_{stay_2}) THEN ($S[t+1]$ is q_2)
R_{33}: IF ($S[t]$ is q_3) AND (ρ is B) AND ($duration_3$ is T_{stay_3}) THEN ($S[t+1]$ is q_3)
R_{44}: IF ($S[t]$ is q_4) AND (((ρ is S) OR (ρ is M)) OR (θ' is D)) AND ($duration_4$ is T_{stay_4}) THEN ($S[t+1]$ is q_4)
R_{12}: IF ($S[t]$ is q_1) AND (((ρ is S) OR (ρ is M)) OR (θ' is D)) AND ($duration_1$ is $T_{change_{12}}$) THEN ($S[t+1]$ is q_2)
R_{23}: IF ($S[t]$ is q_2) AND (ρ is B) AND ($duration_2$ is $T_{change_{23}}$) THEN ($S[t+1]$ is q_3)
R_{34}: IF ($S[t]$ is q_3) AND (((ρ is S) OR (ρ is M)) OR (θ' is D)) AND ($duration_3$ is $T_{change_{34}}$) THEN ($S[t+1]$ is q_4)
R_{41}: IF ($S[t]$ is q_4) AND (ρ is B) AND ($duration_4$ is $T_{change_{41}}$) THEN ($S[t+1]$ is q_1)
R_{01}: IF ($S[t]$ is q_0) AND (ρ is B) AND ($duration_1$ is $T_{change_{41}}$) THEN ($S[t+1]$ is q_1)
R_{02}: IF ($S[t]$ is q_0) AND (((ρ is S) OR (ρ is M)) OR (θ' is D)) AND ($duration_2$ is $T_{change_{12}}$) THEN ($S[t+1]$ is q_2)
R_{03}: IF ($S[t]$ is q_0) AND (ρ is B) AND ($duration_3$ is $T_{change_{23}}$) THEN ($S[t+1]$ is q_3)
R_{04}: IF ($S[t]$ is q_0) AND (((ρ is S) OR (ρ is M)) OR (θ' is D)) AND ($duration_4$ is $T_{change_{34}}$) THEN ($S[t+1]$ is q_4)
R_{i0}: ELSE ($S[t+1]$ is q_0)

where $S[t]$ is the gait state in t, $S[t+1]$ is the gait state in the next sampling time, and $duration_i$ is the membership degree of d_i according to the linguistic labels previously defined.

T produces linguistic expressions that can be adapted depending on the situation. When the signal is in the states q_1, q_2, q_3 or q_4 the template is the following: *"The gait is in the* {*Double limb support*$_1$ | *Right limb single support* | *Double limb support*$_2$ | *Left limb single support*} *state"*. However, when the signal is in the state q_0, that is, the model is not able to recognize the input signal properly, the system reports the following expression: *"The model*

Table 1
Matching degrees obtained for 15 people (in %), depending on the load (0–5) and leg (L/R).

Load	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	P15	Average
0	99	96	98	72	96	98	99	98	99	97	65	84	97	92	90	92
1R	97	97	97	74	91	97	99	93	95	82	67	62	88	83	85	87.1
1L	98	98	99	89	89	99	97	89	96	76	84	64	98	79	88	89.5
2R	95	97	96	81	96	96	97	86	97	75	62	62	99	67	83	85.9
2L	96	99	99	86	97	98	88	78	98	68	81	63	87	76	84	86.5
3R	91	86	97	81	96	99	74	85	98	52	65	57	86	39	86	79.5
3L	92	80	98	71	97	96	91	78	98	60	70	54	93	32	69	78.6
4R	94	75	96	84	95	98	90	61	97	60	78	63	89	23	74	78.5
4L	91	63	98	89	97	91	84	74	97	54	84	55	94	32	83	79.1
5R	90	73	94	88	95	96	61	62	97	63	82	62	92	39	77	78.1
5L	80	60	96	88	98	89	78	73	98	58	82	52	97	37	87	78.2
Deviation	5.27	14.64	1.51	6.91	2.75	3.24	12.04	11.88	1.10	13.62	8.72	8.52	4.69	24.95	6.42	5.3

cannot explain the current situation". If we analyze the rule or rules that make the system to be in the state q_0 we could understand why the signal is uninterpretable. For example, if the signal state goes from q_2 to q_0 because does not fulfill the temporal conditions, the template could report the following sentence: *"The current state of the signal is uninterpretable because the gait has been in right limb single support much longer than it is allowed and the conditions to be in double limb support are not fulfilled."*

4.4. Matching degree of the gait ($2PM_{\mu_U}$)

- U is the input $2CP_{FFSM}$.
- y is the output $2CP_{\mu_U}$, which describes the matching degree of the analyzed gait with respect to the model. $A = (Very\ low\ (VL), Low\ (L), Medium\ (M), High\ (H), Very\ high\ (VH))$.
- g is the aggregation function with calculates the μ_U of the gait for the recording time. With the "uninterpretable" state (q_0), we can define the μ_U of the output based on its activation degrees ($w_{q_0}[t]$) along the time. Such μ_U is defined as shown in Eq. (4). The linguistic labels associated to each expression are represented with trapezoidal membership functions defined as follows: $\{VL\ (-\infty, -\infty, 0.6, 0.65), L\ (0.6, 0.65, 0.7, 0.75), M\ (0.7, 0.75, 0.8, 0.85), H\ (0.8, 0.85, 0.9, 0.95), VH\ (0.9, 0.95, \infty, \infty)\}$.
- T produces linguistic expressions as follows: *"The degree with which the gait suits the model is {very low | low | medium | high | very high}"*.

5. Experimentation

The model of human gait used in this case study has been carefully tested against other models, as result of an intense research to develop a robust and efficient human gait model. For example, in [50] we studied the definition of rules and modal points of linguistic labels by means of genetic algorithms, comparing our results with other methods, such as neural networks and autoregressive models.

The advantages of using the concept of Computational Perception of uninterpretable data (q_0) were checked with the experimentation described below. In Table 1, we show how the matching degree of the model decreases as the gait is more different from the natural one. We show how using this new type of computational perception we can complement the linguistic description of the human gait with a sentence about the degree of matching achieved by the model.

In order to verify the effectiveness and robustness of the developed application, we performed an experiment based on the use of different loads attached to the user's ankles to provoke an imbalanced gait. The smartphone was placed in the waist and centered in the front of the person, measuring the tri-axial accelerations. This experiment involved 15 people, men and women, aged between 25 and 55 years. The procedure consisted of progressively placing loads in the ankles, up to 5 kilograms, walking an approximately distance of 30 meters. We asked the volunteers of the experiment to walk at a regular comfortable speed. Each load was placed alternately in one and another leg, never both at once, in order to cause imbalance between the steps of the left and right legs.

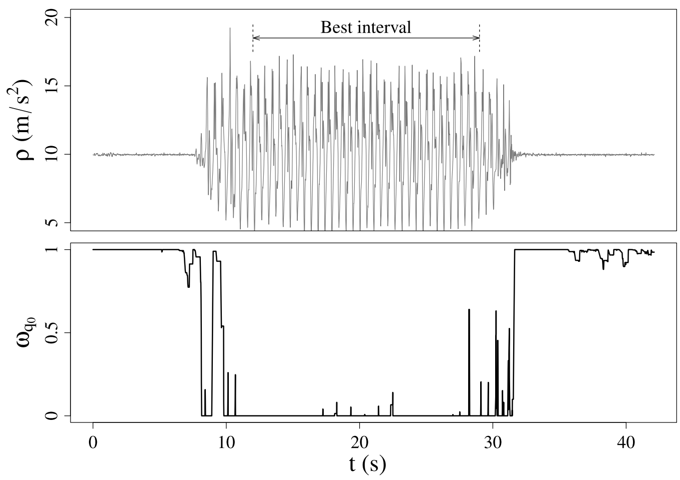


Fig. 13. Original accelerations module (ρ) and the best interval obtained after w_{q_0} analysis.

5.1. Analysis of q_0

The incorporation of the "uninterpretable" state (q_0) within the FFSM allows us to tackle two relevant aspects in the human gait analysis:

- To identify the best time interval to analyze the input signal, that is, to analyze the input data and to see which region is more suitable for gait analysis. Fig. 13 shows the acceleration module (ρ) during one of the performed tests. The evolution of the state q_0 is drawn at the bottom of this picture. Those areas where w_{q_0} is high or close to one are automatically removed for gait analysis. These periods correspond to the times between start/stop data acquisition instants and walking time, i.e., periods when the subject was not walking.
- Once the analysis is done, we can calculate the μ_U of the gait, deciding if either the input data quality is good enough or if it is needed to repeat the data acquisition. Indeed, thanks to q_0, we can complement the linguistic description of the signal with a linguistic expression that remarks the degree with which our model recognizes the phenomenon and, therefore, the reliability of the obtained results.

5.2. Discussion of experimental results

Table 1 shows the matching degrees (μ_U) obtained from the people involved in the experiment (P1 to P15) and the applied loads (from no load ("0") to five kilograms attached in the left ("5L") or right ("5R") leg). The values are expressed in percentages and they indicate how the matching degrees decrease as load increases, so the gait is more imbalanced.

With high loads the user presents a type of lameness that varies depending on the user. As consequence, the obtained gait differs from a "normal" gait and we can identify this variation by observing the value of μ_U. The more distorted the gait is, the higher difference is observed with respect to a reference gait. As expected, the effect of loads on the walking style differs among different people.

We calculated two statistical parameters to analyze in detail the evolution of μ_U. On the one hand, the last column of the table shows the average of μ_U for every load and leg. This column confirms the main hypothesis of the experiment, that is the assumption that matching degrees decrease as load increases. In average, the matching degree of the gait is lower the further the signal is with respect to the "normal" gait used to create the model. Thus, with no load in the legs the average matching degree is 92%, meanwhile with 5 kilograms attached to one leg this value decreases to approximately 78%.

On the other hand, the last row of the table shows the standard deviation obtained for the results of each one of the people. The analysis of people P2, P10 and P14 is specially interesting because μ_U goes from high values with no load to really low values when the loads correspond to four and five kilograms.

We can see how, generally, the trend of the matching degree is decreasing but in some people gaits (P4, P11 and P12), already in absence of load, the signal does not fit well with the designed model. In these cases, the matching degree at the start of the experiment is *Medium*. Under the light of these results, we believe that the initial amplitude and temporal parameters are not the most appropriate ones to model the gait of all the subjects. To solve this problem, Section 6 explains how we can dynamically tune both amplitude and temporal parameters.

In order to generate a linguistic description, we could identify relevant characteristics and categorize different pathologies of the human gait. For example, if we could extract some information about the gait of user P11 with no load attached in his/her legs ("0"), we should start the linguistic report with the following sentence: *"With a matching degree of* 0.65 *over* 1*, we can infer that..."*. To include this information in the final report allows us to express the uncertainty resulting from deficiencies in data analysis or data by themselves.

6. Dynamic tuning of amplitude and temporal parameters

Since the human gait is a quasi-periodic phenomenon which does not produce a perfectly constant and repeatable signal in time, the amplitude and temporal parameters may vary easily. These variations may arise from the comparison between the gait of two different individuals, for different footwear used in tests conducted by the same user, different walking surfaces, etc. In the previous experiment we found out that, even when a person tries to walk uniformly, there are changes in walking speed and applied force.

These variations must be taken into account when we make a robust design that recognizes the gait states, regardless of the amplitude and period of the signal. For this purpose we have developed an automatic tuning of amplitude and temporal parameters, allowing the model to accommodate by itself these slight variations.

Amplitude and temporal parameters are initialized with values that represent a typical gait. The amplitude of the signal is calculated over the accelerations module ($1PM_\rho$) whose linguistic labels were defined in Section 4.1. Temporal parameters were defined in Section 4.3 by the linguistic labels T_{stay_i} and $T_{change_{ij}}$. In the following subsections, we describe how to recalculate dynamically these parameters by considering a temporal sliding window.

6.1. Tuning of temporal parameters

According to the FFSM states recognition, we can calculate the temporal "center of mass" $\bar{t}_i$ of the accelerations module in the state q_i as follows:

$$\bar{t}_i = \frac{\sum_{t=t_0}^{t_f} t \cdot \rho[t] \cdot w_{q_i}[t]}{\sum_{t=t_0}^{t_f} \rho[t] \cdot w_{q_i}[t]} \tag{9}$$

where t_0 is the time instant when the state q_i starts, t_f is the time instant when the state q_i finishes, $\rho[t]$ is the accelerations module at time instant t, and $w_{q_i}[t]$ is the activation degree of the state q_i at time instant t.

With the calculation of this parameter $\bar{t}_i$ we can recalculate dynamically the period K, as $K_i = \bar{t}_i - \bar{t}_{i'}$, where $\bar{t}_{i'}$ is the "center of mass" of the accelerations module in the previously recognized state $q_{i'}$. As a result, we automatically tune the temporal constraints since the linguistic labels T_{stay_i} and $T_{change_{ij}}$ correspond to a percentage of the estimated period. (See Fig. 14.)

6.2. Tuning of amplitude parameters

The way to tune the amplitude constraints consists of modifying the linguistic labels *Small*, *Medium* and *Big* that define the $1CP_\rho$. These labels are defined as follows: $\{$*Small* $(-\infty, -\infty, \bar{\rho} - \sigma_{\bar{\rho}}, \bar{\rho} - 0.25 \cdot \sigma_{\bar{\rho}})$, *Medium* $(\bar{\rho} - \sigma_{\bar{\rho}}, \bar{\rho} - 0.25 \cdot \sigma_{\bar{\rho}}, \bar{\rho}, \bar{\rho} + 0.5 \cdot \sigma_{\bar{\rho}})$, *Big* $(\bar{\rho}, \bar{\rho} + 0.5 \cdot \sigma_{\bar{\rho}}, \infty, \infty)\}$, where $\bar{\rho}$ is dynamically calculated as the average acceleration module and $\sigma_{\bar{\rho}}$ is its standard deviation in the temporal sliding window.

As the dynamic tuning proceeds, the initial parameters are tuned taking into account the requirements of the gait, making the process more robust. To prevent the system from being too sensitive to disturbances, we have introduced a filter which prevents the parameters abruptly change. The value of the inertia parameter (α) can vary depending on the analysis of requirements. Here, we established an inertia parameter equal to 0.8. For example, the period is dynamically adapted as follows: $K = \alpha K_{i'} + (1 - \alpha) K_i$.

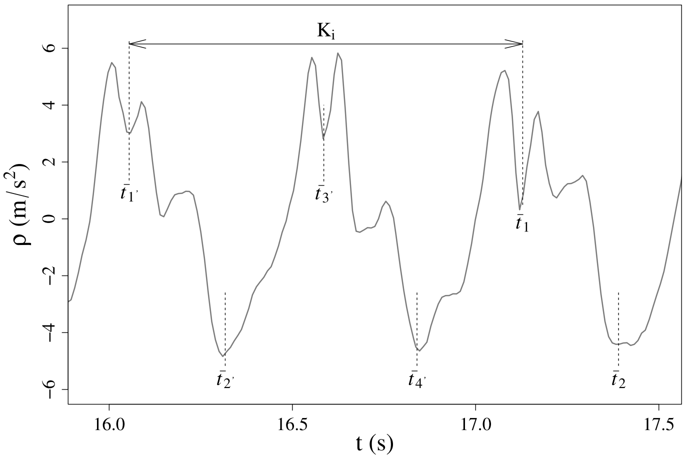


Fig. 14. Center of mass of each state and accelerations module (ρ) minus gravity during the human gait cycle.

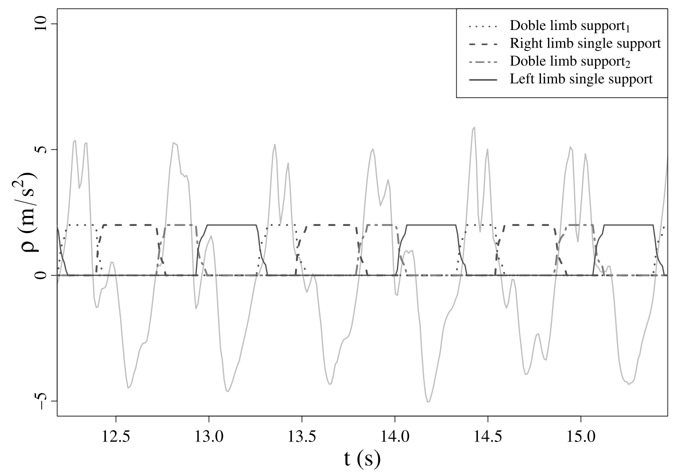


Fig. 15. System output without loads in the legs.

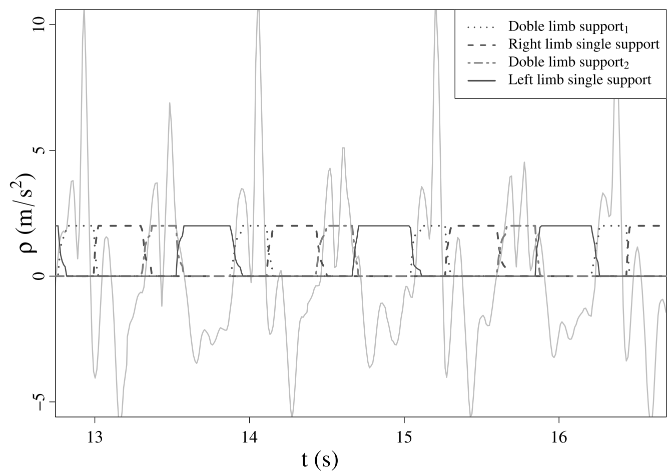


Fig. 16. System output with five kg in the right leg.

7. Experimentation results after dynamical tuning of parameters

Following the experiment introduced in Section 5, here we applied the enhanced model, with dynamic tuning of parameters, to the same set of data. The aim was to check how, despite of changing the gait considerably, the new model recognized correctly the gait phases.

Figs. 15 and 16 show how the accelerations module (ρ) varies in tests performed with different loads. In Fig. 15, we can see how ρ presents a regular evolution. However, in Fig. 16, when the user carries five kilograms in the right leg, we can see how the double support changes significantly from one leg with respect to the other one. In the same way, the swing of the left leg reveals that the force needed to move this leg, and therefore the acceleration produced, is much lower than the force needed to swing the right leg, which has attached five kg.

Table 2 gives the matching degrees for all the tests after including the dynamic tuning in the model. If we compare these results with the reported ones in Table 1, we observe that the dynamic tuning of parameters gets better results for all the performed tests. On the one hand, the average matching degree for each load is bigger than the obtained in Table 1. Here, the average matching degree goes from 98.8% with no load to 96.2% with five kg attached to the left leg. That is, despite having considerable gait variations, the μ_U obtained is always, in average, *Very high*. The standard deviation for each subject decreases, which means that the new model is more independent from these variations. On the other hand, the standard deviation of the average matching degrees was equal to 0.9, which is much lower than the obtained without dynamic tuning, when the result was equal to 5.3. Again, this result demonstrates the effectiveness of the parametric setting driven after the analysis of μ_U, whose definition was based on the activation degrees of the "uninterpretable" state q_0.

Table 2
Matching degrees obtained after dynamical tuning.

Load	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	P15	Average
0	100	100	99	99	100	98	100	99	100	99	99	96	100	93	100	98.8
1R	100	100	96	97	99	96	100	99	100	95	97	96	99	98	89	97.4
1L	100	100	99	98	99	99	98	98	100	96	99	98	99	98	100	98.7
2R	100	100	97	98	99	96	99	87	100	91	98	90	100	99	98	96.8
2L	100	100	99	97	90	98	99	97	99	91	97	96	99	99	100	97.4
3R	100	100	97	98	99	99	99	98	100	95	95	99	97	97	100	98.2
3L	100	100	98	97	91	97	99	97	99	94	99	98	98	98	99	97.6
4R	100	97	96	98	99	99	99	98	100	94	96	98	98	94	99	97.7
4L	99	98	98	94	99	96	99	96	99	94	99	96	98	96	98	97.3
5R	98	95	94	94	99	98	98	98	99	94	94	96	98	90	99	96.3
5L	97	97	96	92	100	91	99	96	99	94	98	93	99	93	99	96.2
Deviation	1.04	1.78	1.60	2.21	3.56	2.32	0.63	3.35	0.52	2.20	1.75	2.57	0.92	2.98	3.17	0.9

Thanks to dynamic tuning, the model is dynamically adapted to each gait, becoming independent of variables such as different surfaces, different footwear, ways of walking, injuries, etc. This is traduced in a better gait phases recognition. It is worth to remark that, once the gait phases identification has been correctly done, we can extract relevant conclusions about the type of walking, possible pathologies and so on, by analyzing the particular parameters of each phase, e.g., symmetry and homogeneity of the gait [38].

Following the example presented in Section 5.2, if we could extract relevant information about the gait of user P11 with no load, now we should start the linguistic report with the following sentence: *"With a matching degree of* 0.99 *over* 1*, we can infer that…"*.

8. Conclusions

This paper presents a contribution to the field of linguistic modeling of complex phenomena and also practical results than can be applied in the field of human gait analysis.

We have extended the definition of Computational Perception by introducing the concept of *uninterpretable data*. Many applications of Fuzzy Logic do not explicitly take into account the meaning of data that cannot be interpreted. In the field of linguistic description of complex phenomena, this new research provides the designer with the possibility of generating more meaningful linguistic reports.

We have used Computational Perceptions of uninterpretable data to design a membership function of a quasi-periodic phenomenon. Using the concept of *uninterpretable data*, we have improved the previous human gait model by making it walking speed and applied force invariant. With this improvement, the model has become more robust and consistent to perform the gait analysis.

This work is part of a long term project aimed to develop computational systems able to generate linguistic models of complex phenomena, where there is still much pending work to do. We believe that the concept of *uninterpretable data* will be useful in the future development of this research line. In this paper we have shown how with a single smartphone, the available results already allow to obtain many relevant features of the human gait.

Experimental results can be used as starting point to model and analyze the details of the human gait. We believe that, perhaps, the new tools will allow us to design thoroughly a model of the human gait able to identify and categorize different pathologies using only a smartphone.

Acknowledgement

This work has been funded by the Spanish Government (MICINN) under project TIN2011-29827-C02-01.

References

[1] M.A. Halliday, C. Matthiessen, Construing experience through meaning: a language-based approach to cognition, Comput. Linguist. 27 (1999).

[2] L.A. Zadeh, Outline of a new approach to the analysis of complex systems and decision processes, IEEE Trans. Syst. Man Cybern. 3 (1973) 28–44.
[3] J.L. Castro, Fuzzy logic controllers are universal approximators, IEEE Trans. Syst. Man Cybern. 25 (1995) 629–635.
[4] L.I. Kuncheva, Fuzzy Classifier Design, Springer-Verlag, London, UK, 2000.
[5] H. Ishibuchi, T. Nakashima, M. Nii, Classification and Modeling with Linguistic Information Granules: Advanced Approaches to Linguistic Data Mining, Adv. Inf. Process., Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2005.
[6] J. Kacprzyk, M. Fedrizzi, Fuzzy Regression Analysis, Stud. Fuzziness Soft Comput., Physica-Verlag, 1992.
[7] P. Pulkkinen, H. Koivisto, A dynamically constrained multiobjective genetic fuzzy system for regression problems, IEEE Trans. Fuzzy Syst. 18 (2010) 161–177.
[8] D. Driankov, H. Hellendoorn, M. Reinfrank, An Introduction to Fuzzy Control, Springer-Verlag, Inc., New York, NY, USA, 1996.
[9] J. Zhang, P. Shi, Y. Xia, Robust adaptive sliding-mode control for fuzzy systems with mismatched uncertainties, IEEE Trans. Fuzzy Syst. 18 (2010) 700–711.
[10] A. Bardossy, L. Duckstein, Fuzzy Rule-Based Modeling with Applications to Geophysical, Biological, and Engineering Systems, CRC Press, Inc., Boca Raton, FL, USA, 1995.
[11] D. Driankov, H. Hellendoorn (Eds.), Fuzzy Model Identification: Selected Approaches, Springer, New York, NY, USA, 1997.
[12] J. Kacprzyk, R. Yager, S. Zadrozny, A fuzzy logic based approach to linguistic summaries of databases, Int. J. Appl. Math. Comput. Sci. (2000) 813–834.
[13] J. Kacprzyk, S. Zadrozny, Computing with words is an implementable paradigm: Fuzzy queries, linguistic data summaries and natural language generation, IEEE Trans. Fuzzy Syst. 18 (2010) 461–472.
[14] J.M. Alonso, L. Magdalena, Special issue on interpretable fuzzy systems, Inf. Sci. 181 (2011) 4331–4339.
[15] W. Pedrycz, Fuzzy modelling: fundamentals, construction and evaluation, Fuzzy Sets Syst. 41 (1991) 1–15.
[16] S. Guillaume, L. Magdalena, Expert guided integration of induced knowledge into a fuzzy knowledge base, Soft Comput. 10 (2006) 773–784.
[17] S. Guillaume, Designing fuzzy inference systems from data: an interpretability-oriented review, IEEE Trans. Fuzzy Syst. 9 (2001) 426–443.
[18] J.M. Alonso, Interpretable fuzzy systems modeling with cooperation between expert and induced knowledge, Ph.D. thesis, Technical University of Madrid, 2007, http://oa.upm.es/588/.
[19] C. Mencar, A.M. Fanelli, Interpretability constraints for fuzzy information granulation, Inf. Sci. 178 (2008) 4585–4618.
[20] C.S. Burrus, R.A. Gopinath, H. Guo, Introduction to Wavelets and Wavelet Transforms: A Primer, Prentice Hall, Upper Saddle River, NJ, 1998, vol. 23.
[21] L.R. Rabiner, A tutorial on hidden Markov models and selected applications in speech recognition, Proc. IEEE 77 (1989) 257–286.
[22] C.M. Bishop, Neural Networks for Pattern Recognition, 1995.
[23] B.D. Ripley, Pattern Recognition and Neural Networks, Cambridge University Press, 2008.
[24] L.A. Zadeh, From computing with numbers to computing with words – from manipulation of measurements to manipulation of perceptions, IEEE Trans. Circuits Syst. I, Fundam. Theory Appl. 45 (1) (1999) 105–119.
[25] L.A. Zadeh, A new direction in AI: Toward a computational theory of perceptions, AI Mag. 22 (1) (2001) 73–84.
[26] L.A. Zadeh, Toward human level machine intelligence – is it achievable? The need for a paradigm shift, IEEE Comput. Intell. Mag. 1 (1) (2008) 1–26.
[27] L. Eciolaza, M. Pereira-Farina, G. Trivino, Automatic linguistic reporting in driving simulation environments, Appl. Soft Comput. (2012).
[28] A. Alvarez-Alvarez, D. Sanchez-Valdes, G. Trivino, A. Sánchez, P.D. Suárez, Automatic linguistic report about the traffic evolution in roads, Expert Syst. Appl. 39 (2012) 11293–11302.
[29] D. Sanchez-Valdes, A. Alvarez-Alvarez, G. Trivino, Linguistic description about circular structures of the Mars' surface, Appl. Soft Comput. 13 (2013) 4738–4749.
[30] L. Arguelles, G. Trivino, i-struve: Automatic linguistic descriptions of visual double stars, Eng. Appl. Artif. Intell. 26 (9) (2013) 2083–2092, http://dx.doi.org/10.1016/j.engappai.2013.05.005.
[31] M.V. Lawson, Finite Automata, Chapman & Hall/CRC, 2003.
[32] J. Davis, Finite State Machine Datapath Design, Optimization, and Implementation (Synthesis Lectures on Digital Circuits and Systems), Morgan and Claypool Publishers, 2008.
[33] E.S. Santos, Maximin automata, Inf. Control 13 (1968) 363–377.
[34] J.N. Mordeson, D.S. Malik, Fuzzy Automata and Languages: Theory and Applications, Comput. Math. Ser., Chapman & Hall/CRC, 2002.
[35] M. Ying, A formal model of computing with words, IEEE Trans. Fuzzy Syst. 10 (2002) 640–652.
[36] Y. Cao, M. Ying, G. Chen, Retraction and generalized extension of computing with words, IEEE Trans. Fuzzy Syst. 15 (2007) 1238–1250.
[37] G. Trivino, A. Alvarez-Alvarez, G. Bailador, Application of the computational theory of perceptions to human gait pattern recognition, Pattern Recognit. 43 (2010) 2572–2581.
[38] A. Alvarez-Alvarez, G. Trivino, Linguistic description of the human gait quality, Eng. Appl. Artif. Intell. 26 (2013) 13–23.
[39] G. Bailador, G. Trivino, Pattern recognition using temporal fuzzy automata, Fuzzy Sets Syst. 161 (2010) 37–55.
[40] D. Sanchez-Valdes, L. Eciolaza, G. Trivino, Linguistic description of human activity based on mobile phone's accelerometers, in: Ambient Assisted Living and Home Care, Springer, 2012, pp. 346–353.
[41] A. Alvarez-Alvarez, J.M. Alonso, G. Trivino, Human activity recognition in indoor environments by means of fusing information extracted from intensity of wifi signal and accelerations, Inf. Sci. (2013) 162–182.
[42] G.J. Klir, B. Yuan, Fuzzy Sets and Fuzzy Logic, Prentice Hall, New Jersey, 1995.
[43] J.B. Yang, J. Liu, J. Wang, H. Sii, H. Wang, Belief rule-base inference methodology using the evidential reasoning approach—RIMER, IEEE Trans. Syst. Man Cybern., Part A, Syst. Hum. (2006) 266–285.
[44] M. Webster online dictionary, http://www.merriam-webster.com, 2013.

[45] E.H. Ruspini, A new approach to clustering, Inf. Control 15 (1969) 22–32.
[46] K. Popper, J. Eccles, The Self and Its Brain, Springer-Verlag, 1977.
[47] K. Ogata, State Space Analysis of Control Systems, Prentice Hall, Englewood Cliffs, 1967.
[48] L.A. Zadeh, Fuzzy sets, Inf. Control 8 (1965) 338–353.
[49] G. Bailador, G. Trivino, A. van der Heide, Fuzzy sets of quasiperiodic signals, in: XIV Spanish Conference on Fuzzy Logic and Technology, ESTYLF, 2008.
[50] A. Alvarez-Alvarez, G. Trivino, O. Cordón, Human gait modeling using a genetic fuzzy finite state machine, IEEE Trans. Fuzzy Syst. 20 (2012).
[51] H. Liu, R.M. Rodríguez, A fuzzy envelope for hesitant fuzzy linguistic term set and its application to multicriteria decision making, Inf. Sci. 258 (0) (2013) 220–238, http://dx.doi.org/10.1016/j.ins.2013.07.027.
[52] C. Alsina, M.J. Frank, B. Schweizer, Associative Functions: Triangular Norms and Copulas, World Scientific Publishing Company, Singapore, 2006.
[53] D.A. Winter, Biomechanics and Motor Control of Human Movement, Wiley–Interscience, 1990.
[54] D.T.H. Lai, R.K. Begg, M. Palaniswami, Computational intelligence in gait research: a perspective on current applications and future challenges, IEEE Trans. Inf. Technol. Biomed. 13 (2009) 687–702.
[55] S.P. Moustakidis, J.B. Theocharis, G. Giakas, A fuzzy decision tree-based SVM classifier for assessing osteoarthritis severity using ground reaction force measurements, Med. Eng. Phys. 32 (2010) 1145–1160.
[56] R. Begg, D.T.H. Lai, M. Palaniswami, Computational Intelligence in Biomedical Engineering, CRC Press, Inc., Boca Raton, FL, USA, 2007.
[57] J. Perry, Gait Analysis: Normal and Pathological Function, 1st edition, Slack Incorporated, 1992.

5.3. Walking pattern classification using a granular linguistic analysis

D. Sanchez-Valdes, A. Alvarez-Alvarez, and G. Trivino. "Walking pattern classification using a granular linguistic analysis". *Applied Soft Computing*, 2015. In press. DOI: 10.1016/j.asoc.2015.04.036

Walking pattern classification using a granular linguistic analysis

Daniel Sanchez-Valdes[a,*], Alberto Alvarez-Alvarez[a], Gracian Trivino[a]

[a]European Centre for Soft Computing,
Edificio de Investigación, c/ Gonzalo Gutiérrez Quirós s/n
33600 Mieres, Asturias, Spain

Abstract

Classifying walking patterns helps the diagnosis of health status, disease progression and the effect of interventions. In this paper, we develop previous research on human gait to extract a meaningful set of parameters that allow us to design a highly interpretable system capable of identifying different gait styles with linguistic fuzzy if-then rules. The model easily discriminates among five different walking patterns, namely: normal walk, on tiptoes, dragging left limb, dragging right limb, and dragging both limbs. We have carried out a complete experimentation to test the performance of the extracted parameters to correctly classify these five chosen gait styles.

Keywords: Walking pattern classification, human gait model, linguistic modeling

1. Introduction

Walking is a learned motor behavior and the simplest, more common and effective mean of locomotion that humans use to transport by only using the body. Among other moderate types of physical activity, walking has strong benefits on a number of aspects of human health.

Gait is a complex task that requires precise coordination of the neural and musculoskeletal system. The measurements from a small number of cycles are representative of the gait pattern of a person. Gait analysis consists of studying the biomechanics of this human movement. The characteristics of our steps are influenced by the shape, position and function of our nervous, muscular and skeletal structures. Therefore, identifying walking patterns can help in the diagnosis of health status, disease progression and the effect of interventions (surgery, rehabilitation or medication), among other medical situations [1, 2, 3].

The importance of detecting dragging limb or walking on tiptoes patterns has been highlighted by several authors [4, 5], due to the fact that a low heel or toe clearance can lead to tripping, which is accounted for up to 50% of all falls. Nevertheless, there is a lack of suitable tools that allow, both users and experts, to measure the relevant characteristics of human gait in an ubiquitous, simple and economic way.

One of the most common approaches to analyze human gait from a kinematic perspective consists of placing accelerometers in the user's body. In [6], authors provide a walking pattern identification (level walking, upstairs, and downstairs) based on an accelerometer placed in the user's ankle. The use of other sensors has also been studied in [7], where authors identify bilateral heel-strike events in data from an inertial measurement unit worn on the waist. Nevertheless, most of these solutions need specific and sophisticated sensors that require the collaboration of experts to be correctly positioned in certain parts of the body or footwear.

*Corresponding author
Email addresses: daniel.sanchezv@softcomputing.es (Daniel Sanchez-Valdes), alberto.alvarez@softcomputing.es (Alberto Alvarez-Alvarez), gracian.trivino@softcomputing.es (Gracian Trivino)

Preprint submitted to Applied Soft Computing *February 10, 2015*

An alternative to specialized sensors is the use of current smartphones. They contain diverse and powerful sensors, e.g., GPS, microphones, cameras, compasses and accelerometers, opening up exciting areas for research in data mining and human-computer interaction. These devices are routinely carried by millions of users and exploit a great range of communications capabilities, integrated hardware and software features.

The accelerations captured during walking can be identified as quasi-periodic signals, i.e., signals with repetitive temporal patterns that include some variations in period and amplitude. Other examples of this type of phenomena are electrocardiograms and vibrations of musical instruments. Popular approaches to deal with quasi-periodic signals vary from Wavelets transform [8], classical curve fitting methodologies, Hidden Markov Models [9] and Neural Networks [10, 11]. Fuzzy Logic (FL) is widely used for linguistic concept modeling and system identification, thanks to its capability for dealing with problems with imprecise information. Thus, FL has been successfully applied in classification, regression and control, achieving a good interpretability accuracy trade-off [12, 13].

In previous research [14, 15, 16], we used Fuzzy Finite State Machines (FFSMs) within the Granular Linguistic Model of a Phenomenon (GLMP) to produce a versatile and interpretable model of the human gait. For the shake of completeness, in this paper, we provide a brief description of these two concepts.

The main structure of this paper is organized as follows:

- In Section 3, we extend our previous human gait model, providing more detailed information about this quasi-periodic phenomenon. We describe the internal details of the FFSM used to recognize the gait phases, achieving a highly interpretable, understandable and efficient model. It is characterized for having linguistic variables and rules that allow to understand the biomedical issues that take part in the gait movement at different levels of detail.

- In Section 4, we use this model to calculate a set of parameters that characterize relevant aspects of the gait. The search and selection of this set of parameters is essential to perform the subsequent discrimination of walking patterns. These parameters are the result of a thorough research based on the relationship between the accelerations and gait phases, and have proven to be a good choice to discriminate the walking patterns.

- In Section 5, the obtained parameters are introduced in a fuzzy rule-based classifier that allows us to discriminate among five different walking patterns, namely: normal walk, on tiptoes, dragging left limb, dragging right limb, and dragging both limbs.

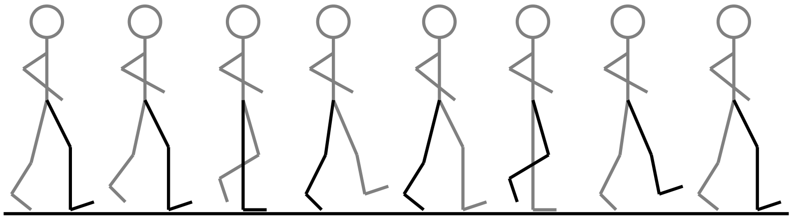
- Finally, in Section 6, we present the complete experimentation carried out to test the performance of this proposal. Ten different subjects were asked to perform the five different walking patterns at self-selected walking speed. The obtained parameters were able to recognize among the five different walking patterns with an overall accuracy of 84%.

2. Preliminary concepts

The core of our technology is the GLMP. The main element of this structure is known as Computational Perception (CP), which is based on the concept of linguistic variable developed by Zadeh [17, 18, 19]. CPs are computational models of units of information (granules) about the phenomenon to be modeled, i.e., CPs correspond to perceptions of specific parts of the phenomenon at certain granularity degree. A CP is a tuple (A, W) described as follows:

A is a vector of linguistic expressions $(a_0, a_1, a_2, \ldots, a_n)$ that represent the whole linguistic domain of the CP. The components of A are defined by the designer in accordance to the most suitable sentences from the typically used ones in the application domain of language. These linguistic expressions will be used to generate linguistic reports about the monitored phenomenon.

W is a vector of validity degrees $(w_0, w_1, w_2, \ldots, w_n)$, where $w_i \in [0, 1]$, assigned to each linguistic expression a_i. In the application context, w_i represents the suitability of a_i to describe the current perception. Since the designer chooses the components of A trying to cover all the possible values of CP, the total



<table>
<tr><td></td><td>Left swing phase (40%)</td><td colspan="3">Left stance phase (60%)</td></tr>
<tr><td colspan="3">Right stance phase (60%)</td><td>Right swing phase (40%)</td><td></td></tr>
<tr><td>Double limb Support (DLS$_1$) (10-12%)</td><td>Right limb single support (RLSS) (36-40%)</td><td>Double limb Support (DLS$_2$) (10-12%)</td><td>Left limb single support (LLSS) (36-40%)</td><td>Double limb Support (DLS$_1$) (10-12%)</td></tr>
</table>

Figure 1: States and events of a gait cycle.

validity should be distributed among all linguistic labels. Therefore, typically, the components of A are associated to fuzzy sets forming strong fuzzy partitions [20] in the universe of discourse of CP, i.e., $\sum_{i=0}^{n} w_i = 1$.

A GLMP consists of a network of Perception Mappings (PMs), which are the elements used to create and aggregate CPs. Each PM receives a set of input CPs and transmits upwards an output CP. In this network, each CP covers specific aspects of the phenomenon with certain granularity degree.

3. Granular Linguistic Model of the human gait

In this paper, we describe a type of GLMP that allows us to model the relevant features of quasi-periodic phenomena evolving in time [21], applied to the human gait.

The gait cycle can be divided into two successive events: the stance phase and the swing phase. By convention, the gait cycle starts when a foot makes contact with the ground. The stance phase then lasts until the same foot is lifted off the ground, at which time the swing phase starts. These two phases can be appreciated in Fig. 1.

Normally, the stance phase represents the 60% of the total gait cycle and the swing phase represents the remaining 40%. These phases can be different among individuals but, when normalized to a percentage of the gait cycle, they maintain close similarity, indicating the absence of disorders [22].

In previous works in this research line, we have built a model that linguistically describes the human gait with different granularity levels. In [16], we presented a FFSM that models CPs to recognize the human gait, building a model that is dynamically self-adapted to users and linguistically describes the human gait. We used this FFSM to describe the following states of the human gait: *Uninterpretable (q_0), Double limb support$_1$ (q_1), Right limb single support (q_2), Double limb support$_2$ (q_3)* and *Left limb single support (q_4)*. Fig. 2 shows the designed model for the linguistic description of the human gait.

Here, we analyze the internal working of this FFSM in detail, increasing the granularity degree of the linguistic descriptions. This detailed analysis allows designers to obtain relevant information about amplitude and temporal changes during gait modeling. In addition, we have extended our previous GLM of the human gait by analyzing the medio-lateral acceleration sign, a new component that will be thoroughly explained in Section 3.3.

Our model differs significantly from others, e.g., based on machine learning techniques, because we use a linguistic model to represent the subjective designer's perceptions of the human gait process. This model is easily understood and does not require high computational cost. Nevertheless, we could use an automatic

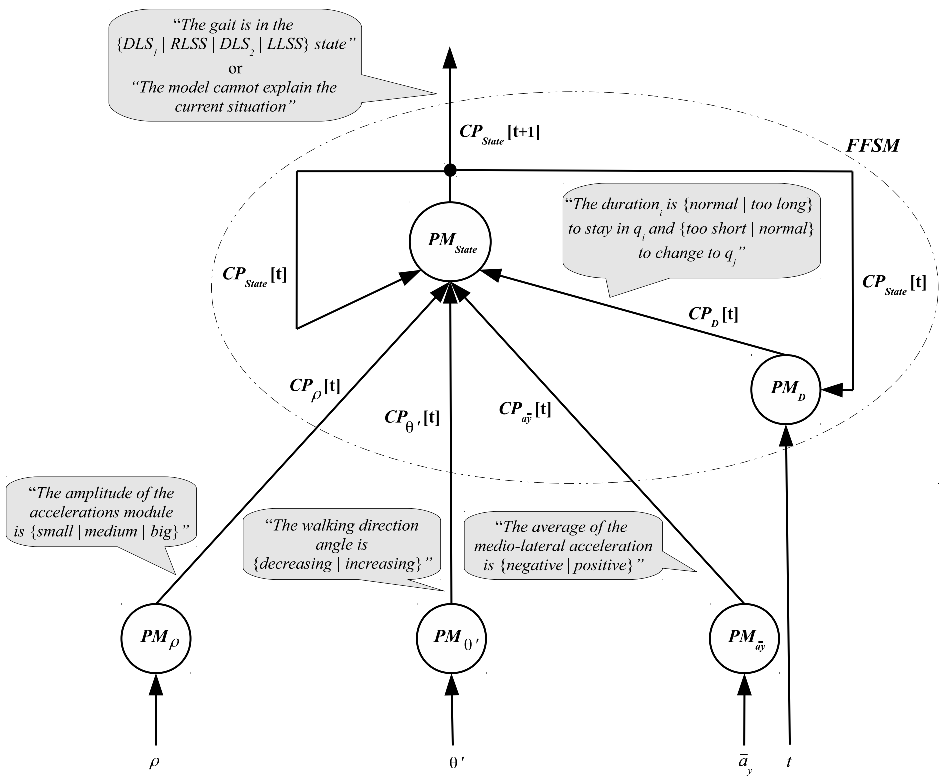


Figure 2: GLMP for the linguistic description of the human gait phases.

machine learning technique to tune the elements of the FFSM as explained in [15], where the interested reader can find a more detailed description of the FFSM paradigm and its applications.

Since the human gait is a quasi-periodic phenomenon, the amplitude and temporal parameters may vary easily due to the comparison between the gait of different individuals, different footwear or different walking surfaces, among others. Moreover, even when a person tries to walk uniformly, there are changes in walking speed and applied force. Our model is automatically self-adapted, recognizing the gait states regardless these variations in the amplitude and period of the signal.

The following subsections briefly describe each component included in this extended GLMP version of the human gait.

3.1. Amplitude of the accelerations module (PM_ρ)

In the case of this PM_ρ, the input corresponds to the accelerations produced during the human gait with a smartphone placed in the waist and centered in the front of the person. This measure provides us the superior-inferior acceleration a_x, the medio-lateral acceleration a_y, and the antero-posterior acceleration a_z at each time instant t (Fig. 3).

The PM_ρ is a tuple (U, y, g, T), where each component is explained as follows:

U are the numerical values of the accelerations module (ρ). The accelerations module is calculated as $\rho = \sqrt{a_x^2 + a_y^2 + a_z^2}$. The possible lack of accuracy when placing the smartphone centered in waist is solved since ρ is an invariant variable respect to the smartphone orientation.

y is the output CP_ρ, $y = (A_y, W_y)$, which describes the possible values of the acceleration module using the vector $A_y =$ (*Small (S), Medium (M), Big (B)*).

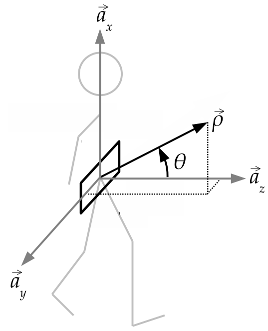


Figure 3: Graphical representation of the smartphone location and provided measures.

g is the output function $W_y = g(\rho)$, where ρ corresponds to the numerical input data. The validity degrees of the output CP are obtained by means of uniformly distributed trapezoidal membership functions (MFs) forming a strong fuzzy partition. The linguistic labels associated to each NL expression are defined by their vertices as follows: $\{S\ (-\infty, -\infty, \bar{\rho}-\sigma_\rho, \bar{\rho}-0.25\cdot\sigma_\rho)$, $M\ (\bar{\rho}-\sigma_\rho, \bar{\rho}-0.25\cdot\sigma_\rho, \bar{\rho}, \bar{\rho}+0.5\cdot\sigma_\rho)$, $B\ (\bar{\rho}, \bar{\rho}+0.5\cdot\sigma_\rho, \infty, \infty)\}$, where $\bar{\rho}$ is the average acceleration module and σ_ρ is the standard deviation, which are dynamically calculated in a temporal sliding window. The interested reader is referred to [16] for more information about this automatic adjustment.

T is a text generation algorithm that allows generating the sentences in A_y. Here, T produces the following linguistic expressions: *"The amplitude of the accelerations module is {small | medium | big}"*.

3.2. Trend of the walking direction angle ($PM_{\theta'}$)

U are the numerical values of the variation of the walking direction angle (θ), where θ is the angle between the accelerations vector ($\vec{\rho}$), and the walking direction represented by the antero-posterior accelerations ($\vec{a}_z$) (Fig. 3). With these numerical values we calculate the derivative of θ (θ'). The use of the derivative of this signal makes the possible accuracy errors smaller than using the absolute values of the angle

y is the output $CP_{\theta'}$, which describes the possible values of θ' using the vector A_y = (*Decreasing (D), Increasing (I)*). It is mainly used to help us to determine when the double limb support phases finish.

g is the output function similar to the corresponding one of PM_ρ. Here, the two linguistic labels associated to each NL expression are represented with trapezoidal MFs defined by their vertices as follows: $\{D\ (-\infty, -\infty, -0.1, 0.1),\ I\ (-0.1, 0.1, \infty, \infty)\}$.

T produces the following linguistic expressions: *"The walking direction angle is {decreasing | increasing}"*.

3.3. Medio-lateral acceleration sign ($PM_{\bar{a}_y}$)

U are the numerical values of the sign of the average medio-lateral acceleration (a_y) during a temporal moving window of 0.5 seconds.

y is the output $CP_{\bar{a}_y}$, which describes the possible values of the sign of $\bar{a}_y$ using the vector A_y = (*Negative (N), Positive (P)*). Since, a priori, the movement of both limbs are similar, this CP is mainly used to distinguish between the two legs.

g is the output function that calculates the validity degrees of the output $CP_{\bar{a}_y}$. Here, the two linguistic labels associated to each NL expression are represented with trapezoidal MFs defined by their vertices as follows: $\{N\ (-\infty, -\infty, -0.05, 0.05),\ P\ (-0.05, 0.05, \infty, \infty)\}$.

T produces the following linguistic expressions: *"The average of the medio-lateral acceleration is {negative | positive}"*.

3.4. State duration (PM_D)

U is the vector of input CPs, $U = (u_1, u_2, \ldots, u_m)$, where u_i are tuples (A_i, W_i) and m the number of input CPs. Here, U are the input CP_{State} and the timespan t. Each state q_i has an associated duration d_i, which is numerically calculated using t.

y is the output CP_D which describes the duration of each state q_i. The possible linguistic expressions to describe this perception are contained in the vector: A_y = (*too short to change, normal to change, normal to stay, too long to stay*).

g is the output function that calculates the validity degrees of CP_D. When the validity degree of the state q_i (w_{q_i}) takes the value zero, its duration d_i takes value zero. However, when w_{q_i} takes values greater than zero, its duration d_i increases its value according to the time, measuring the duration of the state.

Based on the duration d_i of each state, we define two temporal conditions described as follows:

- The time to stay is the maximum time that the signal is allowed to stay in state q_i. We define two linguistic labels that describe this perception: {*normal to stay, too long to stay*}.
- The time to change is the minimum time that the signal must be in state q_i before changing to other state q_j. We define other two linguistic labels that describe this perception: {*too short to change, normal to change*}.

According to the dynamic adjustment of these temporal conditions, the vertices that define their MFs are expressed in terms of the gait period K, that initially takes value 1 (1 gait cycle per second). The way of calculating this period will be explained in Section 4 (Eq. 7).

In this application, due to the similarities between the states q_1 and q_3, their temporal constraints are the same and they are represented by trapezoidal MFs defined by their vertices as follows:

- Time to stay: {*normal* $(-\infty, -\infty, 0.2 \cdot K, 0.3 \cdot K)$, *too long* $(0.2 \cdot K, 0.3 \cdot K, \infty, \infty)$}.
- Time to change: {*too short* $(-\infty, -\infty, 0.05 \cdot K, 0.15 \cdot K)$, *normal* $(0.05 \cdot K, 0.15 \cdot K, \infty, \infty)$}.

In the same way, the states q_2 and q_4 share the temporal constraints, which are defined as:

- Time to stay: {*normal* $(-\infty, -\infty, 0.35 \cdot K, 0.45 \cdot K)$, *too long* $(0.35 \cdot K, 0.45 \cdot K, \infty, \infty)$}.
- Time to change: {*too short* $(-\infty, -\infty, 0.2 \cdot K, 0.3 \cdot K)$, *normal* $(0.2 \cdot K, 0.3 \cdot K, \infty, \infty)$}.

T produces the following linguistic expressions: *"The duration is {normal | too long} to stay in q_i and {too short | normal} to change to q_j"*, with $i \in \{1, 2, 3, 4\}$.

3.5. Gait states (PM_{State})

Here, we show how to build a FFSM using the aggregation of CPs. The set of components of this PM is explained as follows:

U is the vector of input CPs: $\{CP_\rho$, $CP_{\theta'}$, $CP_{\bar{a}_y}$, CP_D, $CP_{State}\}$.

y is the output CP_{State}, which describes the possible states of the gait as the states of a FFSM. A_y = (*Uninterpretable (q_0), Double limb support$_1$ (q_1), Right limb single support (q_2), Double limb support$_2$ (q_3), Left limb single support (q_4)*). Note that the output CP_{State} is also an input of the vector U, that feeds the PM_{State} and calculates the next state at time instant $[t + 1]$ taking into account the current state at time instant $[t]$.

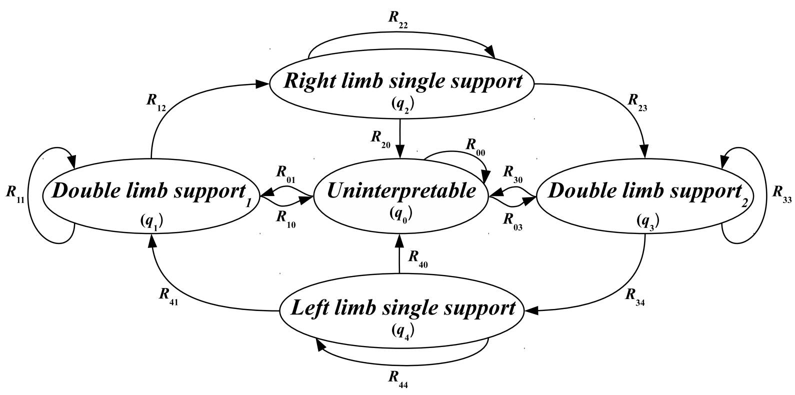


Figure 4: States diagram.

g is the aggregation function $W_y = g(W_1, W_2, \ldots, W_m)$, where W_i are the vectors of validity degrees of the m input CPs. The set of amplitude conditions (CP_ρ, $CP_{\theta'}$, $CP_{\bar{a}_y}$), the temporal condition (CP_D) and the current state of the signal (CP_{State}) implement this FFSM model.

The aggregation function g implements a set of fuzzy rules according to the states diagram showed in Fig. 4. We have rules to remain in each state q_i (R_{ii}) and rules to change from state q_i to state q_j (R_{ij}). The set of rules has been designed using the experts' knowledge in human gait as follows:

R_{11}: IF ($State[t]$ is q_1) AND (ρ is B) AND (d_1 is *normal to stay in* q_1)
THEN ($State[t+1]$ is q_1)
R_{22}: IF ($State[t]$ is q_2) AND (((ρ is S) OR (ρ is M)) OR (θ' is D)) AND (d_2 is *normal to stay in* q_2)
THEN ($State[t+1]$ is q_2)
R_{33}: IF ($State[t]$ is q_3) AND (ρ is B) AND (d_3 is *normal to stay in* q_3)
THEN ($State[t+1]$ is q_3)
R_{44}: IF ($State[t]$ is q_4) AND (((ρ is S) OR (ρ is M)) OR (θ' is D)) AND (d_4 is *normal to stay in* q_4)
THEN ($State[t+1]$ is q_4)
R_{12}: IF ($State[t]$ is q_1) AND (((ρ is S) OR (ρ is M)) OR (θ' is D)) AND (d_1 is *normal to change to* q_2)
THEN ($State[t+1]$ is q_2)
R_{23}: IF ($State[t]$ is q_2) AND (ρ is B) AND (d_2 is *normal to change to* q_3)
THEN ($State[t+1]$ is q_3)
R_{34}: IF ($State[t]$ is q_3) AND (((ρ is S) OR (ρ is M)) OR (θ' is D)) AND (d_3 is *normal to change to* q_4)
THEN ($State[t+1]$ is q_4)
R_{41}: IF ($State[t]$ is q_4) AND (ρ is B) AND (d_4 is *normal to change to* q_1)
THEN ($State[t+1]$ is q_1)
R_{01}: IF ($State[t]$ is q_0) AND (ρ is B) AND ($\bar{a}_y$ is *positive*)
THEN ($State[t+1]$ is q_1)
R_{03}: IF ($State[t]$ is q_0) AND (ρ is B) AND ($\bar{a}_y$ is *negative*)
THEN ($State[t+1]$ is q_3)
R_{i0}: ELSE ($State[t+1]$ is q_0)

where $State[t]$ is the gait phase at time instant t, ρ is the amplitude of the accelerations module, d_i is the duration of the state q_i and $State[t+1]$ is the gait phase in the next time instant. The final values of this gait phase $State[t+1]$ are calculated as a weighted average of the individual rules, where the weight of each rule R_{ij} corresponds to its firing degree τ_{ij}. This firing degree is calculated using the product for the AND operator and the bounded sum of Łukasiewicz [23] for the OR operator. We chose the state *Uninterpretable* (q_0) as initial state of the phenomenon, having a validity degree

$w_{q_0} = 1$. The validity degree of the state q_0 is obtained by means of Eq. 1:

$$w_{q_0}[t+1] = \begin{cases} 1 - \sum_{i=0}^{n}\sum_{j=1}^{n} \tau_{ij} & \text{if} \quad \sum_{i=0}^{n}\sum_{j=1}^{n} \tau_{ij} \leq 1 \\ 0 & \text{if} \quad 1 < \sum_{i=0}^{n}\sum_{j=1}^{n} \tau_{ij} \end{cases} \tag{1}$$

The validity degree of the rest of states is calculated by means of Eq. 2:

$$w_{q_j}[t+1] = \begin{cases} \sum_{i=0}^{n} \tau_{ij} & \text{if} \quad \sum_{i=0}^{n}\sum_{j=1}^{n} \tau_{ij} \leq 1 \\ \frac{\sum_{i=0}^{n} \tau_{ij}}{\sum_{i=0}^{n}\sum_{j=1}^{n} \tau_{ij}} & \text{if} \quad 1 < \sum_{i=0}^{n}\sum_{j=1}^{n} \tau_{ij} \end{cases} \tag{2}$$

T produces linguistic expressions that can be adapted depending on the situation. When the signal is in the states q_1, q_2, q_3 or q_4 the template is the following: *"The gait is in the {Double limb support$_1$ | Right limb single support | Double limb support$_2$ | Left limb single support} phase"*. However, when the signal is in the state q_0, i.e., the model is not able to recognize the input signal properly, the system reports the following expression: *"The model cannot explain the current situation"*.

4. Set of characteristic parameters of the human gait

This section describes how to calculate those parameters that can discriminate among different walking patterns. First, we describe how to calculate them at the level of each state or gait phase. Second, we divide each state in order to obtain additional parameters at the *initial*, *middle* and *final* level of each phase. Then, we also calculate other ones associated to the homogeneity and symmetry of the gait. Finally, we calculate some additional features that combine parameters from the same gait phase of the two different limbs.

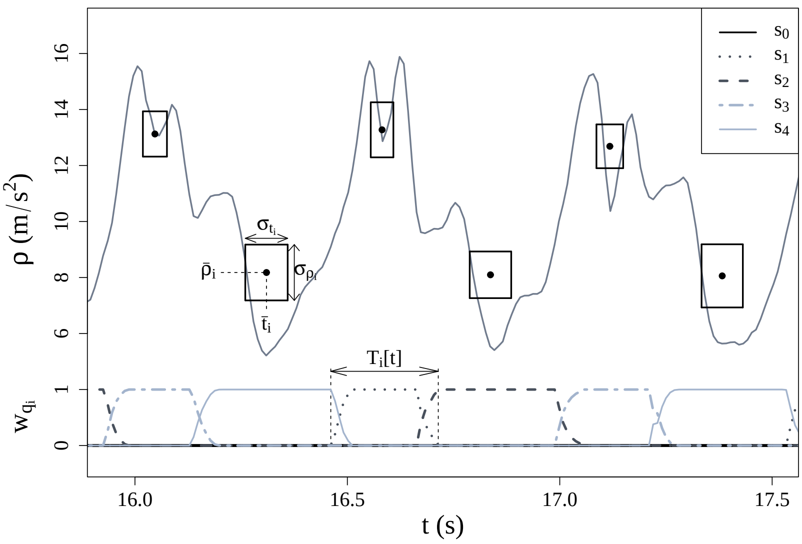


Figure 5: Characteristic rectangles, acceleration module (ρ) and states definition during the human gait.

4.1. Parameters of each phase

Based on the phases identification presented in Section 3, we can obtain a set of characteristic parameters of each state q_i. Each phase can be represented by a rectangle of variable dimensions, as shown in Fig. 5. Note that this graphical representation allows us to acquire a first visual perception of the signal characteristics. The parameters that define these rectangles are calculated as follows:

- T_i: Is the duration of state q_i within a gait cycle. This duration corresponds to the time that the validity degree of the state q_i is greater than zero, i.e., $w_{q_i} > 0$.
- $\bar{t_i}$: The horizontal coordinate of the center of each rectangle is the temporal "center of mass" of the accelerations module in the state q_i:

$$\bar{t_i} = \frac{\sum_{t=0}^{T_i} t \cdot \rho[t] \cdot w_{q_i}[t]}{\sum_{t=0}^{T_i} \rho[t] \cdot w_{q_i}[t]} \tag{3}$$

- $\bar{\rho_i}$: The vertical coordinate of the center of each rectangle is the average of the accelerations module during the state q_i:

$$\bar{\rho_i} = \frac{\sum_{t=0}^{T_i} \rho[t] \cdot w_{q_i}[t]}{\sum_{t=0}^{T_i} w_{q_i}[t]} \tag{4}$$

- σ_{t_i}: The width of each rectangle is the standard deviation of the temporal distribution of the accelerations module during the state q_i:

$$\sigma_{t_i} = \sqrt{\frac{\sum_{t=0}^{T_i} (t - \bar{t_i})^2 \cdot \rho[t] \cdot w_{q_i}[t]}{\sum_{t=0}^{T_i} \rho[t] \cdot w_{q_i}[t]}} \tag{5}$$

- σ_{ρ_i}: The height of each rectangle is the standard deviation of the accelerations module during the state q_i:

$$\sigma_{\rho_i} = \sqrt{\frac{\sum_{t=0}^{T_i} (\rho[t] - \bar{\rho_i})^2 \cdot w_{q_i}[t]}{\sum_{t=0}^{T_i} w_{q_i}[t]}} \tag{6}$$

where $\rho[t]$ is the accelerations module at time instant t, and $w_{q_i}[t]$ is the validity degree of the state q_i at time instant t.

The period of the gait is dynamically calculated using the parameter $\bar{t_i}$. When each state q_i finishes, a new period $K_i[n]$ is obtained as $\bar{t_i}[n] - \bar{t_i}[n-1]$, being n the current gait cycle and $n-1$ the previous one. To prevent the system from being too sensitive to disturbances, we have introduced an inertia parameter ($\alpha = 0.8$) which avoids that the period of the signal $K[t]$ changes abruptly. Thus, the period is calculated as follows:

$$K[n+1] = \alpha K[n] + (1-\alpha)K_i[n] \tag{7}$$

where $K[n]$ is the previous calculated period.

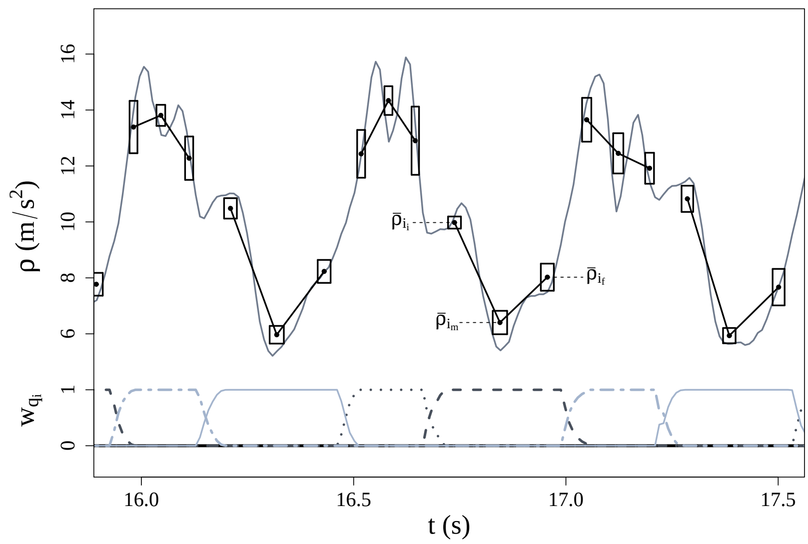


Figure 6: Characteristic rectangles of each subphase, acceleration module (ρ) and states definition during the human gait.

4.2. Parameters of each subphase

Apart from obtaining the characteristic rectangles of each phase, and according to the typical decomposition in gait analysis, we divide each state q_i into three equal parts, namely, *initial*, *middle* and *final* subphases. In the double limb support phases, these subphases go from the heel strike and foot-flat of one limb, to the toe-off of the another one. In the single limb support phases, these subphases correspond to the initial swing, midswing and terminal swing. Then, we calculate, as explained at the beginning of this section, the representative rectangles of each of these subphases, as shown in Fig. 6.

Here, we are interested in the relationship among rectangles of the same state q_i. Specifically, we analyze the evolution of the vertical coordinate of the center of each rectangle for the three representative rectangles of each state ($\bar{\rho}_{i_i}$, $\bar{\rho}_{i_m}$ and $\bar{\rho}_{i_f}$). As we will explain in Section 5, the relation among these vertical coordinates will be different according to the gait style of users. Fig. 7 shows the typical relation between these three parameters according to the walking pattern. As we can visually appreciate, there is a characteristic pattern depending on whether the user is walking normal, on tiptoes or dragging the limb.

In order to measure how these vertical coordinates vary, we calculate two additional parameters for each state, namely A_i and Ω_i that calculate the difference among these values as follows:

$$\begin{aligned} A_i &= \frac{\sum_{n=1}^{N} \bar{\rho}_{i_m}[n] - \bar{\rho}_{i_i}[n]}{N} \\ \Omega_i &= \frac{\sum_{n=1}^{N} \bar{\rho}_{i_f}[n] - \bar{\rho}_{i_m}[n]}{N} \end{aligned} \tag{8}$$

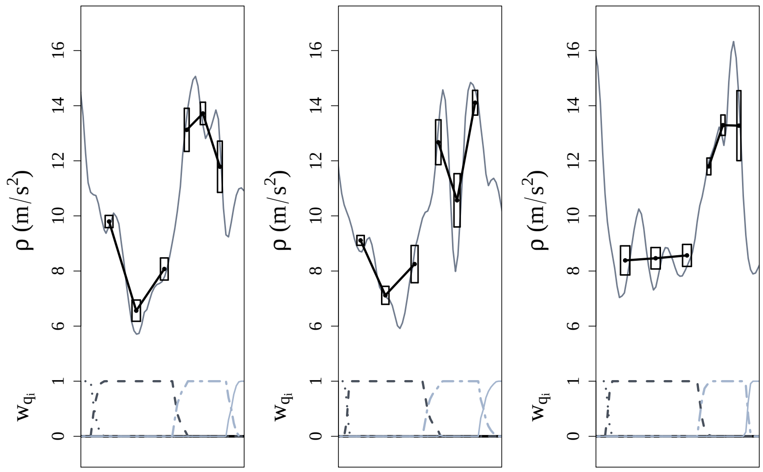
where, n is the gait cycle number and N is the total number of gait cycles in the recorded data.

4.3. Homogeneity and Symmetry

Once the phases and subphases are recognized, we calculate two relevant features of the human gait (homogeneity and symmetry) for three of the characteristic parameters explained in Section 4.1, specifically, for $\bar{\rho}_i$, σ_{ρ_i} and T_i.

The homogeneity of a gait is obtained by comparing the same phase in two consecutive gait cycles. Thus, we can analyze the homogeneity during each state q_i and obtain the parameters $H_{\bar{\rho}_i}$, $H_{\sigma_{\rho_i}}$ and H_{T_i}

The symmetry of a gait is obtained by comparing a phase with its equivalent, i.e., by comparing the swing phases of the left and the right limbs or the two different double limb support phases. Thus, we can



(a) Normal (b) On tiptoes (c) Dragging

Figure 7: Characteristic patterns depending on users' gait style.

analyze the symmetry during the double limb support phases and obtain the parameters $S_{\bar{\rho}_{13}}$, $S_{\sigma_{\rho_{13}}}$ and $S_{T_{13}}$, or during the swing phases and obtain the parameters $S_{\bar{\rho}_{24}}$, $S_{\sigma_{\rho_{24}}}$ and $S_{T_{24}}$.

Both, homogeneities and symmetries of the gait, are calculated using the Jaccard index [24], whose similarity function $J(A, B)$ is defined in Eq. 9:

$$J(A,B) = \begin{cases} 1 & \text{if } A = B = 0 \\ \frac{|A \cap B|}{|A \cup B|} & \text{otherwise} \end{cases} \tag{9}$$

where the intersection $|A \cap B|$ is implemented as $\min(A, B)$ and the union $|A \cup B|$ as $\max(A, B)$.

For example, consider that the vertical coordinate during the left swing phase $\bar{\rho}_2$ has been 7.89 m/s^2 and 7.34 m/s^2 during two consecutive gait cycles $k-1$ and k respectively. Therefore, the homogeneity of these two cycles will be calculated using Eq. 10:

$$H_{\bar{\rho}_2}[k] = J(\bar{\rho}_2[k], \bar{\rho}_2[k-1]) = \frac{\min(\bar{\rho}_2[k], \bar{\rho}_2[k-1])}{\max(\bar{\rho}_2[k], \bar{\rho}_2[k-1])} = \frac{7.34}{7.89} = 0.93 \tag{10}$$

4.4. Additional parameters

To generalize some of the parameters explained above, we have calculated six additional parameters that combine two of them, corresponding to the same gait phase that were calculated for each different limb, in a single one:

- A_{13}: is calculated as the average value between the parameters A_1 and A_3, i.e., the variation of the vertical coordinate between the *middle* ($\bar{\rho}_{i_m}$) and *initial* ($\bar{\rho}_{i_i}$) subphases of the double support phase (q_1 and q_3).

- A_{24}: is calculated as the average value between the parameters A_2 and A_4, i.e., the variation of the vertical coordinate between the *middle* ($\bar{\rho}_{i_m}$) and *initial* ($\bar{\rho}_{i_i}$) subphases of the swing phase (q_2 and q_4).

- Ω_{13}: is calculated as the average value between the parameters Ω_1 and Ω_3, i.e., the variation of the vertical coordinate between the *final* ($\bar{\rho}_{i_f}$) and *middle* ($\bar{\rho}_{i_m}$) subphases of the double support phase (q_1 and q_3).

- Ω_{24}: is calculated as the average value between the parameters Ω_2 and Ω_4, i.e., the variation of the vertical coordinate between the *final* ($\bar{\rho}_{i_f}$) and *middle* ($\bar{\rho}_{i_m}$) subphases of the swing phase (q_2 and q_4).
- T_{31}: is calculated as the average difference between the durations of the double support phases (q_3 and q_1) within a complete gait cycle.
- T_{42}: is calculated as the average difference between the durations of the swing phases (q_4 and q_2) within a complete gait cycle.

5. Walking pattern classification

One limitation of the gait analysis arises from the variability of each walking pattern within the general population. Even people with no neuromusculoskeletal pathology have very different gait patterns. For this reason, experts need a tool that objectively helps them in the diagnosis of health status, disease progression and the effect of interventions in those pathologies that affect the ability to walk properly. Nowadays, to monitor the evolution of some rehabilitation processes experts use, among others, subjective methods that consist of recalling questionnaires. The alternative consists in going to clinical laboratories that analyze movement, or placing specific and sophisticated sensors to provide an objective measure that allows them to monitor the gait evolution of patients. In this sense, there is a lack of suitable tools that measure the relevant characteristics of human gait in a ubiquitous, simple and economic way. In this paper, we propose a system that allows experts to identify five well known walking patterns while they can monitor the evolution of patients along the time by only using a simple smartphone.

5.1. Feature selection

With the aim of distinguish among different walking patterns, we initially tried to capture and identify the relevant gait phases and a big amount of parameters that are representative of these phases. Once this first step was completed, we noticed that these data provide enough information for this identification purpose.

However, since our goal consists of providing a granular linguistic analysis for this identification, we tried to reduce the number of available parameters. Therefore, the second step consisted of applying a feature selection technique in order to extract the most discriminatory among the initial 50 parameters obtained from the gait analysis explained in Section 4. There are many machine learning algorithms that can be used to select the most appropriate attributes for a classification task. In this work, we have tried some of them, e.g., Support Vector Machines as attribute evaluator and a ranker as search method. In order to keep the interpretability of the whole system, we have finally chosen the decision tree algorithm proposed by Quinlan [25]. This algorithm chooses the most promising attribute to split on at each point and should, in theory, never select irrelevant or unhelpful attributes. The attribute space searching was done with a forward selection using a 10-fold cross validation. It was implemented using Weka software [26], a well-known machine learning tool.

Finally, once we get a set of relevant features, we choose those ones that make sense from an interpretable point of view. Moreover, during this third step, we observed that there were some discriminatory parameters that can represented by the single ones explained in Section 4.4. Therefore, we chose five discriminative parameters that were fuzzified by means of linguistic labels whose definition has been obtained from the experimental training data:

- A_{24} is defined by two linguistic labels represented by their vertices as follows:
  $\{$*Negative* $(-\infty, -\infty, -1.8, -1.4)$, *Zero* $(-1.8, -1.4, \infty, \infty)\}$.
- Ω_{13} is defined by two linguistic labels represented by their vertices as follows:
  $\{$*Negative* $(-\infty, -\infty, -1.3, -1)$, *Positive* $(-1.3, -1, \infty, \infty)\}$.
- Ω_{24} is defined by two linguistic labels represented by their vertices as follows:
  $\{$*Zero* $(-\infty, -\infty, 0.8, 1.3)$, *Positive* $(0.8, 1.3, \infty, \infty)\}$.

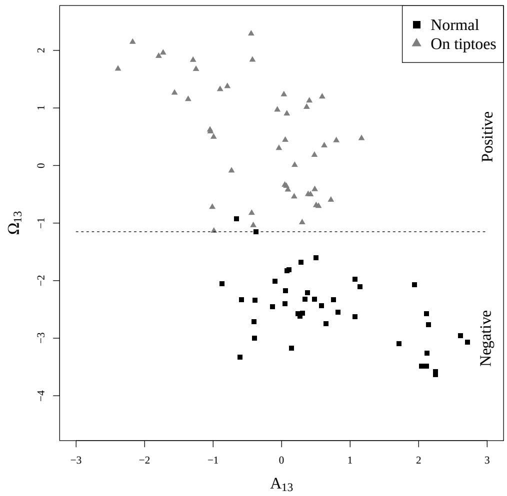


Figure 8: Graphical representation of the parameters A_{13} and Ω_{13} for the double support phases.

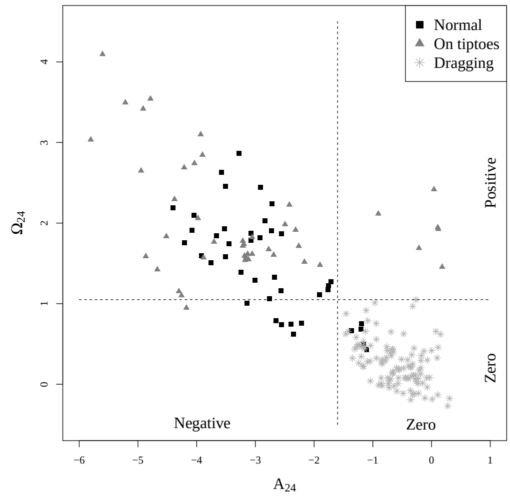


Figure 9: Graphical representation of the parameters A_{24} and Ω_{24} for the swing phases.

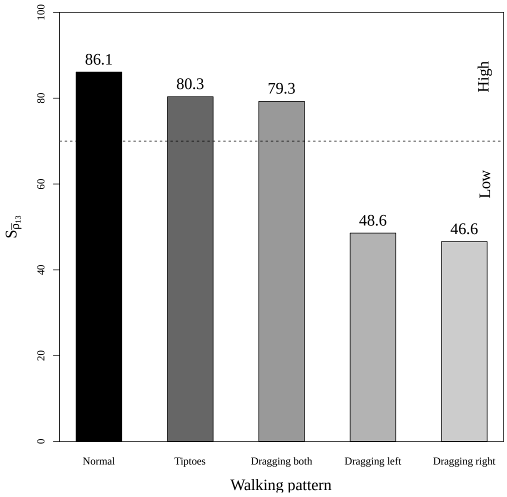


Figure 10: Barplot of the symmetry during the double limb support phase.

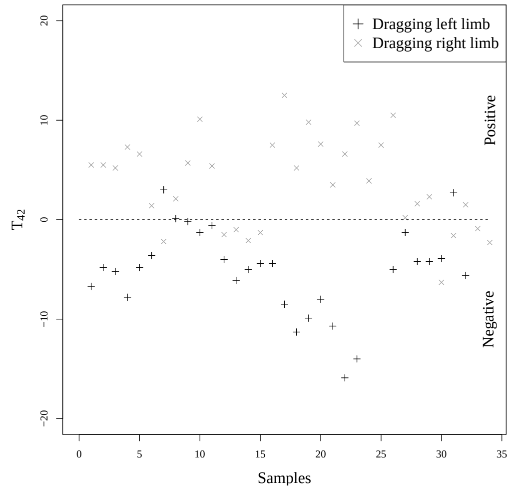


Figure 11: Average difference duration between the swing phases represented by the states q_2 and q_4.

- $S_{\bar{\rho}_{13}}$ is defined by two linguistic labels represented by their vertices as follows: $\{Low\ (-\infty, -\infty, 68, 72),\ High\ (68, 72, \infty, \infty)\}$.
- T_{42} is defined by two linguistic labels represented by their vertices as follows: $\{Negative\ (-\infty, -\infty, -1.5, 1.5),\ Positive\ (-1.5, 1.5, \infty, \infty)\}$.

As we can appreciate in Figs. 8-11, it is easy to visually differentiate among walking patterns using these parameters. These figures represent in dashed lines the cutoffs of the linguistic labels explained before. Fig. 8 shows the use of Ω_{13} to discriminate between "normal" and "on tiptoes" patterns; Fig. 9 shows the use of A_{24} and Ω_{24} to discriminate the "dragging" pattern, does not making distinction among dragging both legs or only one of them; Fig. 10 shows that the symmetry during the double limb support ($S_{\bar{\rho}_{13}}$) is much lower when the user is "dragging one limb" that in the rest of patterns; and, finally, Fig. 11 shows the use of T_{42} to discriminate between "dragging left limb" and "dragging right limb" patterns. In this way, we have designed a highly interpretable model that is able to linguistically describe the reasons that produce a specific walking pattern identification.

5.2. Set of rules

Once we know which the five discriminatory parameters are, we used them to build an expert knowledge fuzzy rule-based classifier that differentiates among the chosen walking patterns. The set of fuzzy rules that represents this fuzzy rule-based classifier is the following:

R_1: IF (A_{24} is *negative*) AND (Ω_{13} is *negative*)
THEN (the user *walks normal*)
R_2: IF (A_{24} is *negative*) AND (Ω_{13} is *positive*)
THEN (the user *walks on tiptoes*)
R_3: IF (A_{24} is *zero*) AND (Ω_{24} is *positive*) AND (Ω_{13} is *negative*)
THEN (the user *walks normal*)
R_4: IF (A_{24} is *zero*) AND (Ω_{24} is *positive*) AND (Ω_{13} is *positive*)
THEN (the user *walks on tiptoes*)
R_5: IF (A_{24} is *zero*) AND (Ω_{24} is *zero*) AND ($S_{\bar{\rho}13}$ is *high*)
THEN (the user *drags both limbs*)
R_6: IF (A_{24} is *zero*) AND (Ω_{24} is *zero*) AND ($S_{\bar{\rho}13}$ is *low*) AND (T_{42} is *negative*)
THEN (the user *drags the left limb*)
R_7: IF (A_{24} is *zero*) AND (Ω_{24} is *zero*) AND ($S_{\bar{\rho}13}$ is *low*) AND (T_{42} is *positive*)
THEN (the user *drags the right limb*)

where the firing degree τ_i of each rule R_i is calculated using the product for the AND operator. The fuzzy rule-based classifier can be graphically represented by the fuzzy decision tree showed in Fig. 12.

Note that the set of fuzzy rules provides a linguistic explanation about the obtained walking pattern. For example, when the system concludes that the user is dragging the right limb, it can be explained analyzing the antecedents of rule R_7:

- A_{24} is *zero*, which means that the initial and middle subphases of the swing phases are similar.
- Ω_{24} is *zero*, which means that the middle and final subphases of the swing phases are similar.
- $S_{\bar{\rho}13}$ is *low*, which means that symmetry during the double supports is low because one limb is being dragged.
- T_{42} is *positive*, which means that the duration of state q_4 (swing phase of the dragged limb) is greater than the duration of state q_2 (left swing phase).

Therefore, the classifier will provide the following message: *"the subject is dragging the right limb, because the duration of the right limb swing phase is greater than the one corresponding to the left limb. Moreover, the symmetry during the double limb support is low due to this dragging right limb"*. For a more detailed description about linguistic description of complex phenomena, we address the interested reader to [27, 28].

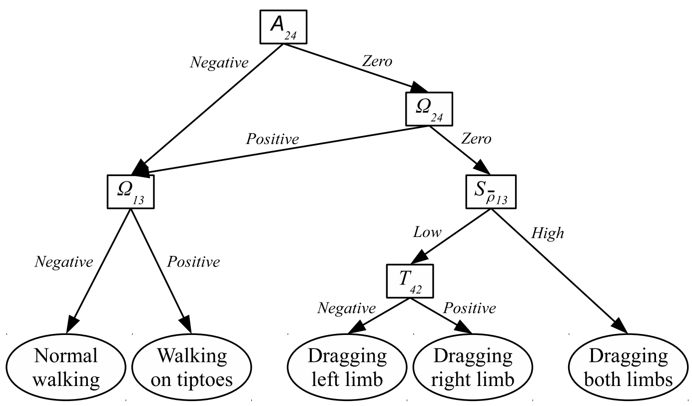


Figure 12: Fuzzy decision tree for the walking patterns identification.

6. Experimentation

This section presents the experiments performed to show the potential and effectiveness of our approach. This experimentation can be considered only the first step for testing the viability and possibilities of a future complete and functional tool. We will need the collaboration of experts in different pathologies to tune the gait model and overcome the analysis of gaits affected by real physical or mental disorders, such as cerebral palsy and Parkinson's disease. The following subsections detail the most relevant aspects of this experimental phase.

6.1. Objective

The objective of the experimentation was to demonstrate the discrimination power of the obtained parameters. With this aim, we divided it into a training and a test stage.

During the training stage, we performed the feature selection to extract the relevant parameters, and tuned the fuzzy membership functions and rules of the fuzzy rule-based classifier as explained in Section 5.

Then, in the test stage, we evaluated the capability of the designed model to correctly identify among the five different walking patterns.

6.2. Experimental layout

In order to evaluate the proposed approach, we collected the acceleration signals of 10 healthy adults, 5 women and 5 men, with ages ranging between 24 and 57 years (with an average age of 34 years).

In order to collect the accelerations produced during walking, we developed an application based in the Android platform that allows the smartphone to acquire and store the accelerations produced during the gait. We have chosen Android-based smartphones because the Android operating system is free, open-source, and it has become a dominant entry in the smartphones marketplace.

We attached the smartphone to a belt, on the waist and centered in the front of the person. The smartphone provides the timestamp, the superior-inferior (a_x), the medio-lateral (a_y), and the antero-posterior (a_z) accelerations with an average acquisition frequency of 100 Hz.

Each person received detailed instructions about how to perform each of the five different walking patterns and was asked to walk a distance of approximately 30 meters on a flat surface at a self-selected walking speed. The whole data collection process was always supervised by the same person, in order to avoid errors when placing the smartphone and with the aim of keeping a repetitive methodology.

The collected data comprises approximately twenty complete gait cycles, in such a way that we were able to automatically extract around fifteen complete gait cycles discarding the first and last steps, which are not very stable. This process was repeated five times for each walking pattern, producing a total of 25 datasets for each person and resulting in a total of 250 different gaits (50 gaits for each walking pattern).

We used 200 gaits (40 of each walking pattern from all of the participants) in the training stage (80% of the available data) to perform the feature selection and tunning of the fuzzy membership functions and rules of the fuzzy rule-based classifier. The remaining 50 gaits (10 of each walking pattern from all of the participants) were used in the test stage (20% of the available data) to evaluate the performance of our proposal.

6.3. Results

Table 1 shows the confusion matrix corresponding to the identification results. We have obtained an average identification accuracy of 84%. The best identification accuracies were obtained with the normal walk (100%), walk on tiptoes (90%), and dragging both limbs (90%) patterns. On the contrary, the worst identification accuracies corresponded to the dragging left and right limb patterns (70% for both of them).

Table 2 provides some statistical measures, obtained for each walking pattern, that inform about the performance of our proposal. The first two columns contain the true positive (TP) and false positive (FP) rates. Third, fourth and fifth columns show the precision, recall and the F-Measure, respectively. Finally, last column shows the area under the Receiver Operating Characteristic (ROC) curve. In addition, last row indicates the average values from the previous measures, giving us an idea about the identification success.

Table 1: Confusion matrix.

		Predicted: Normal	On tiptoes	Dragging both	Dragging left	Dragging right	Accuracy (%)
Actual	Normal	10	0	0	0	0	100
	On tiptoes	0	9	1	0	0	90
	Dragging both	0	0	9	1	0	90
	Dragging left	0	0	3	7	0	70
	Dragging right	0	0	2	1	7	70

Table 2: Detailed average accuracy with the designed fuzzy decision tree.

Walking pattern	**TP**	**FP**	**Precision**	**Recall**	**F-Measure**	**ROC Area**
Normal	1	0	1	1	1	1
On tiptoes	0.9	0	1	0.9	0.947	0.949
Dragging both limbs	0.9	0.15	0.6	0.9	0.72	0.91
Dragging left limb	0.7	0.05	0.778	0.7	0.737	0.831
Dragging right limb	0.7	0	1	0.7	0.824	0.896
Average	0.84	0.04	0.876	0.84	0.846	0.917

6.4. Discussion

It is important to remark that each of the five walking patterns were performed by each of the ten healthy individuals according to their physical characteristics. Therefore, the dragging patterns (left, right or both limbs) may be quite different among each person. That is the reason why the accuracy of these gait styles is smaller than the rest of styles. Moreover, since the gait is a complex process that implies coordination of both limbs, when the person tries to drag only one of the two limbs, the movement of the other one could be affected by this unnatural movement, resulting in a "dragging both limbs" pattern. This fact can be easily checked in Table 1, where we can see how the prediction of "dragging the left limb" was correctly identified the 70% of times, meanwhile the remaining 30% of test data was identified as "dragging both limbs". Something similar occurs for the "dragging right limb pattern", where 20% of test data was identified as a "dragging both limbs" pattern.

Currently, we are working in an ambitious experimental project in collaboration with rehabilitation experts at Valle del Nalón Hospital (Langreo, Asturias) in order to measure pathological gaits from real patients. The goal is to tune the model and explore the possibility of creating practical tools that help experts in the diagnosis and rehabilitation processes.

Table 3: Detailed accuracy comparison with other classification algorithms.

Algorithm	**TP**	**FP**	**Precision**	**Recall**	**F-Measure**	**ROC Area**
Naïve Bayes	0.8	0.05	0.812	0.8	0.799	**0.957**
Multilayer Perceptron	0.82	0.045	0.824	0.82	0.821	0.949
SVM	0.8	0.05	0.871	0.8	0.809	0.932
Our proposal	**0.84**	**0.04**	**0.876**	**0.84**	**0.846**	0.917

6.5. Comparison with other classification algorithms

In order to check the accuracy of the presented algorithm with respect to other classical classification ones, we have compared our results with the ones obtained with Naïve Bayes, Multilayer Perceptron and Support Vector Machines algorithms. Table 3 shows the average values of each statistical measure for each of these classification algorithms in comparison with our proposal results. In bold letters are highlighted the best results by each of these statistical measures.

Although similar, the overall accuracy of the fuzzy decision tree is better (84%) than the ones obtained with the Naïve Bayes (80%), Multilayer Perceptron (82%) and Support Vector Machines (80%) algorithms. Results reveal that in this application case, where not only the overall accuracy is the most important factor, but also the interpretability of the model to provide classification explanations, the designed fuzzy decision tree is a good solution to discriminate among different walking patterns.

7. Conclusions and future works

In this paper, we present a contribution towards the development of new tools for human gait analysis. We describe the internal details of the Fuzzy Finite State Machine used to recognize the gait phases, achieving an interpretable, understandable and efficient model. It allows to understand the biomedical issues that take part in the gait movement at different levels of detail.

Based on the information related to the accelerations and gait phases, we have designed a model that extracts a broad set of parameters used to characterize and to identify five well known walking patterns: normal walk, on tiptoes, dragging left limb, dragging right limb, and dragging both limbs.

This work is part of a long term and ambitious project that will need the collaboration of multidisciplinary experts in human gait to tune the model and analyze real pathological gaits produced by cerebral palsy and Parkinson's disease, among others. This is a first approach aimed at obtaining a functional tool that helps experts in the diagnosis and tracking of several physical and mental disorders.

Experimental results reveal the capability of the model to correctly identify the different gait styles. As an extension of previous research, this paper shows a promising and exciting research line that will deliver useful medical applications in the near future.

Acknowledgment

This work has been funded by the Spanish Government (MICINN) under the project TIN2011-29827-C02-01 and the Principality of Asturias Government under the projects CT13-51 and CT13-52.

References

[1] D. T. H. Lai, R. K. Begg, M. Palaniswami, Computational intelligence in gait research: a perspective on current applications and future challenges, Information Technology in Biomedicine, IEEE Transactions on 13 (5) (2009) 687–702.

[2] S. P. Moustakidis, J. B. Theocharis, G. Giakas, A fuzzy decision tree-based svm classifier for assessing osteoarthritis severity using ground reaction force measurements, Medical Engineering & Physics 32 (10) (2010) 1145–1160.

[3] D. T. H. Lai, M. Palaniswami, R. Begg, Healthcare Sensor Networks: Challenges Toward Practical Implementation, CRC Press, 2011.

[4] B. Mariani, S. Rochat, C. J. Bla, K. Aminian, Heel and toe clearance estimation for gait analysis using wireless inertial sensors., IEEE Transactions on Biomedical Engineering 59 (11-2) (2012) 3162–3168.

[5] W. P. Berg, H. M. Alessio, E. M. Mills, C. Tong, Circumstances and consequences of falls in independent community-dwelling older adults, Age Ageing 26 (4) (1997) 261–8.

[6] J.-S. Wang, C.-W. Lin, Y.-T. C. Yang, Y.-J. Ho, Walking pattern classification and walking distance estimation algorithms using gait phase information., IEEE Transactions on Biomedical Engineering 59 (10) (2012) 2884–2892.

[7] M. Yuwono, S. W. Su, Y. Guo, B. D. Moulton, H. T. Nguyen, Unsupervised nonparametric method for gait analysis using a waist-worn inertial sensor, Applied Soft Computing 14, Part A (0) (2014) 72–80.

[8] C. S. Burrus, R. A. Gopinath, H. Guo, Introduction to wavelets and wavelet transforms: a primer, Vol. 23, Prentice hall Upper Saddle River NJ, 1998.

[9] L. R. Rabiner, A tutorial on hidden Markov models and selected applications in speech recognition, Proceedings of the IEEE 77 (2) (1989) 257–286.

[10] C. M. Bishop, Neural networks for pattern recognition, Clarendon press Oxford, 1995.

[11] B. D. Ripley, Pattern recognition and neural networks, Cambridge University Press, 2008.
[12] A. Bardossy, L. Duckstein, Fuzzy Rule-Based Modeling with Applications to Geophysical, Biological, and Engineering Systems, CRC Press, Inc, Boca Raton, FL, USA, 1995.
[13] D. Driankov, H. Hellendoorn (Eds.), Fuzzy Model Identification: Selected Approaches, Springer, New York, NY, USA, 1997.
[14] G. Trivino, A. Alvarez-Alvarez, G. Bailador, Application of the computational theory of perceptions to human gait pattern recognition, Pattern Recognition 43 (7) (2010) 2572–2581.
[15] A. Alvarez-Alvarez, G. Trivino, O. Cordón, Human gait modeling using a genetic fuzzy finite state machine, Fuzzy Systems, IEEE Transactions on 20 (2) (2012) 205–223.
[16] D. Sanchez-Valdes, G. Trivino, Computational perceptions of uninterpretable data. A case study on the linguistic modeling of human gait as a quasi-periodic phenomenon, Fuzzy Sets and Systems 253 (2014) 101–121.
[17] L. A. Zadeh, The concept of a linguistic variable and its application to approximate reasoning - I, Information Sciences 8 (3) (1975) 199–249.
[18] L. A. Zadeh, The concept of a linguistic variable and its application to approximate reasoning - II, Information Sciences 8 (4) (1975) 301–357.
[19] L. A. Zadeh, The concept of a linguistic variable and its application to approximate reasoning - III, Information Sciences 9 (1) (1975) 43–80.
[20] E. H. Ruspini, A new approach to clustering, Information and Control 15 (1969) 22–32.
[21] D. Sanchez-Valdes, G. Trivino, Increasing the granularity degree in linguistic descriptions of quasi-periodic phenomena, Flexible Query Answering Systems, Springer (2013) 281–292.
[22] J. Perry, Gait Analysis: Normal and Pathological Function, 1st Edition, Slack Incorporated, 1992.
[23] C. Alsina, M. J. Frank, B. Schweizer, Associative Functions: Triangular Norms and Copulas, World Scientific Publishing Company, Singapore, 2006.
[24] L. Hamers, Y. Hemeryck, G. Herweyers, M. Janssen, H. Keters, R. Rousseau, A. Vanhoutte, Similarity measures in scientometric research: the Jaccard index versus Salton's cosine formula, Information Processing & Management 25 (3) (1989) 315–318.
[25] J. R. Quinlan, Induction of decision trees, Machine Learning 1 (1986) 81–106.
[26] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, I. H. Witten, The weka data mining software: an update, ACM SIGKDD explorations newsletter 11 (1) (2009) 10–18.
[27] D. Sanchez-Valdes, A. Alvarez-Alvarez, G. Trivino, Linguistic description about circular structures of the Mars' surface, Applied Soft Computing 13 (12) (2013) 4738–4749.
[28] L. González-Villanuevaa, A. Alvarez-Alvarez, L. Ascari, G. Trivino, A tool for linguistic assessment of rehabilitation exercises, Applied Soft Computing 14 (Part A) (2014) 120–131.

Daniel Sanchez-Valdes received his M.S. degree (2011) in Industrial Engineering from the University of Oviedo (Oviedo, Spain). In June 2012, he received the M.S. degree in Soft Computing and Intelligent Data Analysis jointly organized by the European Centre for Soft Computing and the University of Oviedo. He is currently working toward the Ph.D. degree in Computer Science with the same university. Since September 2011 he has been Predoctoral Researcher at the European Centre for Soft Computing. He has published papers in international journals and peer-reviewed conferences, and he has received the Best Special Session Student Paper Award in the 11th International Conference on Intelligent Systems Design and Applications (Córdoba, Spain, 2011). His main research interests are focused on the generation of linguistic descriptions from non-linguistic data, such as sensor signals or database systems.

Alberto Alvarez-Alvarez received his M.S. degree in Physics (2008) from the Universidad Autónoma de Madrid (Spain). In 2010, he received the M.S. degree in Soft Computing and Intelligent Data Analysis jointly organized by the European Centre for Soft Computing and the University of Oviedo. Finally, in 2012, he received the Ph.D. Degree in Computer Science from the same university. Since September 2008, he has been Predoctoral Researcher at the European Centre for Soft Computing. Since January 2013, he is Postdoctoral Researcher in the same research centre. He has published more than fifteen papers in international journals and peer-review conferences. His main research interests include the use of sensors to obtain and manage data from the physical context, analyze them using Soft Computing techniques and produce accurate models close to natural language. Mr. Alvarez Alvarezs three contributions have received the Best Special Session Student Paper Award in the 11th International Conference on Intelligent Systems Design and Applications (Córdoba, Spain, 2011), the IEEE Computational Intelligence Society 2011 Outstanding Paper Award at the 5th IEEE International Workshop on Genetic and Evolutionary Fuzzy Systems (Paris, France, 2011), and the Best Presentation Award at the International Open Workshop Fuzziness and Medicine (Mieres, Spain, 2011).

Gracian Trivino received his M.S. and Ph.D. degrees in Electrical Engineering from the Polytechnic University of Madrid (Madrid, Spain), in 1982 and 2000, respectively. From 1982 to 1999 he occupied several positions managing R&D projects in industry. His research areas covered during this professional activity include Industrial Control, Nuclear Power Plants Control, Telecommunications, Avionics, and Software Engineering. From 1999 to 2006 he was Associate Professor with the Computer Science School of the Polytechnic University of Madrid. Research topics covered were Intelligent building control, Control of physical processes, Electronic instrumentation, Photonic devices, Development of computational systems, and Architectures for robotics. Since January 2007 he is Principal Researcher with the European Centre for Soft Computing. During these years, he has participated in 20 public funded projects, has written 30 papers in journals, has registered seven patents and has sent more than 50 contributions to peer-reviewed scientific congresses. His current research includes the development of Zadeh's Computational Theory of Perceptions toward linguistic description of complex phenomena.

Capítulo 6

Factor de impacto

El factor de impacto (en inglés *Impact Factor* - IF), es una medida calculada anualmente que refleja el número medio de citas que reciben las revistas de ciencias y ciencias sociales por sus artículos más recientes. En particular, este factor se calcula para aquellas revistas que están indexadas en el *Journal Citation Reports* ®, de Thomson Reuters. Este factor se utiliza frecuentemente como una forma de medir la importancia relativa que tiene una revista dentro de su campo de actuación, siendo más importantes, por lo tanto, aquellas revistas que tienen mayores factores de impacto.

Las siguientes secciones recogen los *Journal Citation Report* (JCR) del año 2013, correspondientes a cada una de las revistas en donde se publicaron los artículos presentados en el Capítulo 5. Aunque alguno de estos artículos se han publicado posteriormente al año 2013, los JCR de los años 2014 y 2015 se comunicarn a mediados de 2015 y 2016, respectivamente.

6.1. JCR de Applied Soft Computing

ISI Web of Knowledge℠

Journal Citation Reports®

2013 JCR Science Edition

Journal: APPLIED SOFT COMPUTING

Journal Impact Factor

Cites in 2013 to items published in:	2012 = 704	Number of items published in:	2012 = 341
	2011 = 1670		2011 = 545
	Sum: 2374		Sum: 886

Calculation: Cites to recent items / Number of recent items = 2374 / 886 = **2.679**

Journal Ranking

For **2013**, the journal **APPLIED SOFT COMPUTING** has an Impact Factor of **2.679**.

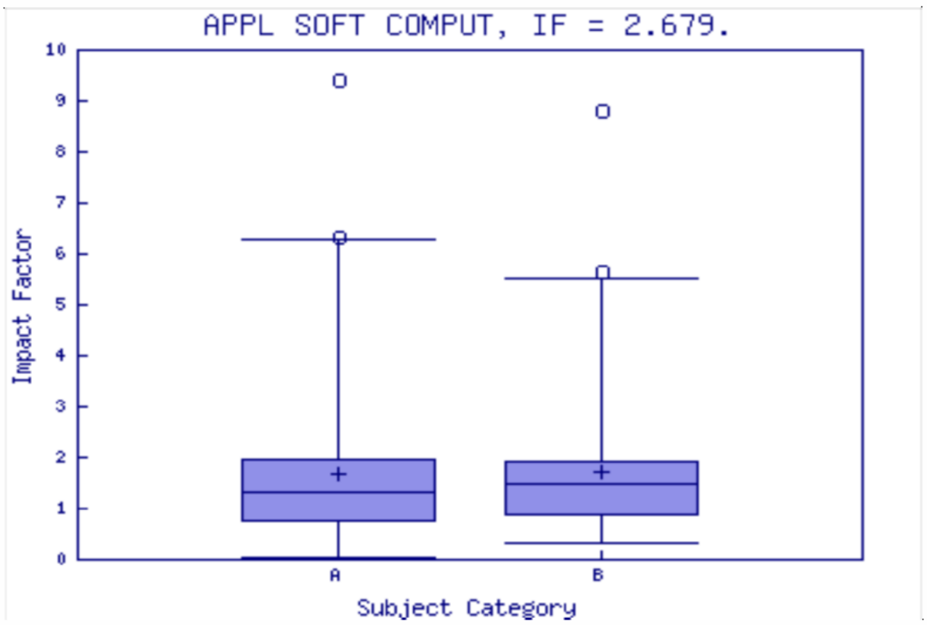
This table shows the ranking of this journal in its subject categories based on Impact Factor.

Category Name	Total Journals in Category	Journal Rank in Category	Quartile in Category
COMPUTER SCIENCE, ARTIFICIAL INTELLIGENCE	121	20	Q1
COMPUTER SCIENCE, INTERDISCIPLINARY APPLICATIONS	102	14	Q1

Category Box Plot

For **2013**, the journal **APPLIED SOFT COMPUTING** has an Impact Factor of **2.679**.

This is a box plot of the subject category or categories to which the journal has been assigned. It provides information about the distribution of journals based on Impact Factor values. It shows median, 25th and 75th percentiles, and the extreme values of the distribution.



Key

A - COMPUTER SCIENCE, ARTIFICIAL INTELLIGENCE

B - COMPUTER SCIENCE, INTERDISCIPLINARY APPLICATIONS

6.2. JCR de Fuzzy Sets and Systems

ISI Web of Knowledge[SM]

Journal Citation Reports®

2013 JCR Science Edition

Journal: FUZZY SETS AND SYSTEMS

Journal Impact Factor

Cites in 2013 to items published in: 2012 = 262 Number of items published in: 2012 = 171
2011 = 381 2011 = 171
Sum: 643 Sum: 342

Calculation: Cites to recent items / Number of recent items = 643 / 342 = **1.880**

Journal Ranking

For **2013**, the journal **FUZZY SETS AND SYSTEMS** has an Impact Factor of **1.880**.

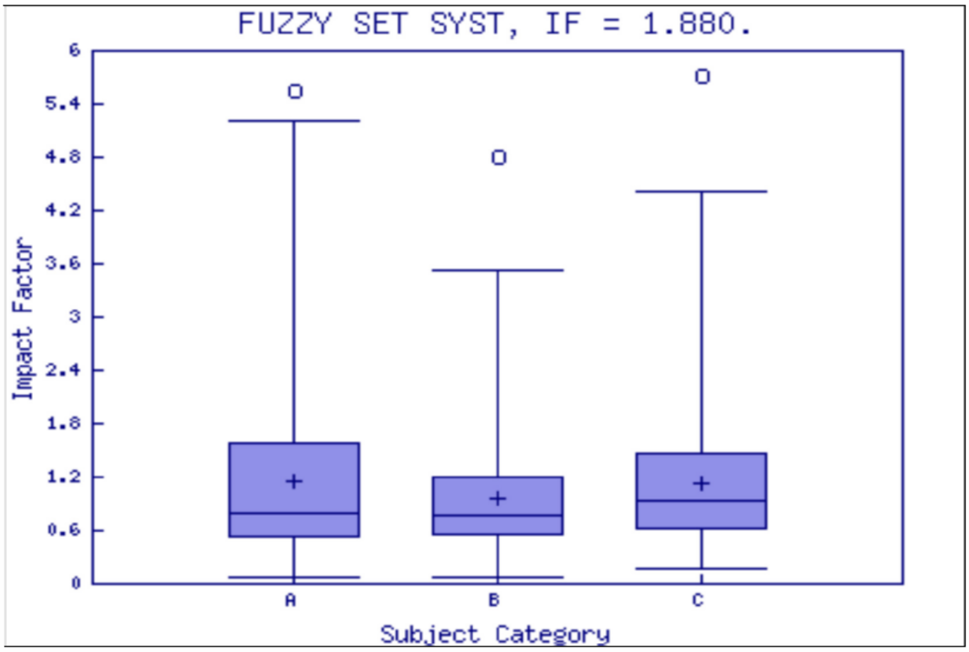
This table shows the ranking of this journal in its subject categories based on Impact Factor.

Category Name	Total Journals in Category	Journal Rank in Category	Quartile in Category
COMPUTER SCIENCE, THEORY & METHODS	102	16	Q1
MATHEMATICS, APPLIED	250	20	Q1
STATISTICS & PROBABILITY	119	14	Q1

Category Box Plot

For **2013**, the journal **FUZZY SETS AND SYSTEMS** has an Impact Factor of **1.880**.

This is a box plot of the subject category or categories to which the journal has been assigned. It provides information about the distribution of journals based on Impact Factor values. It shows median, 25th and 75th percentiles, and the extreme values of the distribution.



Key

A - COMPUTER SCIENCE, THEORY & METHODS

B - MATHEMATICS, APPLIED

C - STATISTICS & PROBABILITY

Capítulo 7

Publicaciones adicionales

Este capítulo contiene cuatro publicaciones que, aunque no se han presentado en el Capítulo 5, están muy relacionadas con los objetivos y líneas de investigación desarrolladas durante la tesis. Se divide en cuatro secciones diferentes, correspondientes a cada artículo junto con su referencia bibliográfica.

7.1. Linguistic Description of Human Activity Based on Mobile Phone's Accelerometers

D. Sanchez-Valdes, L. Eciolaza, and G. Trivino. "Linguistic Description of Human Activity Based on Mobile Phones Accelerometers". *Ambient Assisted Living and Home Care*, pp. 346–353, Springer.

Linguistic Description of Human Activity Based on Mobile Phone's Accelerometers

Daniel Sanchez-Valdes, Luka Eciolaza, and Gracian Trivino

European Centre for Soft Computing (ECSC)
Mieres, Asturias, Spain
{daniel.sanchezv,luka.eciolaza,gracian.trivino}@softcomputing.es

Abstract. Monitoring the physical activity of a person is used for many applications such as medical assistance, personal security, etc. We use the accelerometers embedded in current mobile phones to identify the physical activities and generate periodical linguistic reports which describe relevant information about the activities carried out, their intensities and trends. Based on the Computational Theory of Perceptions, we have developed an application able to generate these linguistic descriptions, presenting a simple demonstration of our contribution in this field.

Keywords: Linguistic summarization, Computing with perceptions.

1 Introduction

Activity recognition has captured the attention of several computer science communities due to its strength in providing personalized support for many different applications and its connection to many different fields of study such as medicine, human-computer interaction, or sociology.

Monitoring the body posture and the physical activity of a person can be useful for applications such as medical assistance, trying to identify falls or abnormalities in the course of daily activity. Until recently, the evaluation of a patient's health outside a doctor's office has been limited to questionnaires which were hindered by memory and patients' subjectivity. New technological devices are now providing information which was previously impossible to attain. Continuous health monitoring of a patient could lead to faster and better diagnosis of problems. It could also be used to gauge patient's progress during rehabilitation where medicines and physical treatments could be better controlled.

Using three axial accelerometers in order to recognize the body posture and activities of a person is a well-known area of study. The analysis of the personal physical activities is used to assess the health of a subject, taking into account that walking gives a good indication of the energy expenditure of an individual. However, so far the use of accelerometers has been limited to a lab setting and could not be used in continuous health monitoring.

In this paper, we analyze the physical activities performed by people along a period of time, typically one day, only by taking acceleration data from their mobile phones. Nowadays common use devices, such as mobile phones, allow

J. Bravo, R. Hervás, and M. Rodríguez (Eds.): IWAAL 2012, LNCS 7657, pp. 346–353, 2012.

acquiring data from a patient in an inexpensive way. The sensors embedded to these devices are very accurate, easy to use and non-intrusive for the patients.

We use natural language (NL) to describe patterns emerging in data by means of linguistic expressions, choosing the most adequate granularity degree in each circumstances, in the same way that humans describe their perceptions. Linguistic Description of Data (LDD) is intended in general for applications in which there is a strong human-machine interaction involving accessing and understanding data. Therefore, it is well suited for activity recognition and description purposes.

Fuzzy logic (FL) [1] and the Computational Theory Perceptions (CTP) [2] are the basic theories of our LDD approach. FL is widely recognized for its ability in linguistic concept modeling. FL evolved to CTP, which provides a framework to implement computational systems with the capacity of computing with the meaning of NL expressions, i.e. with the capacity of computing with imprecise descriptions of the world in a similar way that humans do it. In previous works [3][4][5][6], we have developed the architecture to implement CTP for the linguistic descriptions of complex phenomena.

This paper describes how to apply our approach for linguistic description of complex phenomena to recognize and report the daily physical activities of a person. We correctly identify activities as important as walking, standing, sitting, walking upstairs and walking downstairs. In addition, knowing the quantity of time spent in each activity, we can approximately calculate the user caloric expenditure.

This paper is organized as follows. Section 2 describes the architecture of a system able to create linguistic descriptions of phenomena while section 3 explains how to apply it in the field of activity recognition. Afterwards, section 4 shows the experimentation and validation carried out. Finally, section 5 provides some concluding remarks.

2 Architecture

The main processing modules of this computational system are, namely, the *Data Acquisition (DAQ) module*, the *Validity module*, and the *Expression module*.

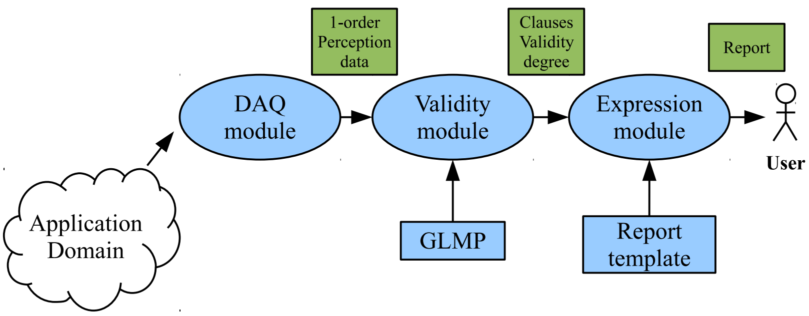


Fig. 1. Main components of the proposed computational system for linguistic description of data

Fig. 1 shows the architecture used for automatic report generation in the application case presented in this paper. The *DAQ module* could include either sensors or access to database information, representing the interface with the physical environment. This module provides the data needed to feed up the *Validity module*, which uses the linguistic model (GLMP) to generate a collection of linguistic sentences of the phenomenon with associated validity degrees.

Finally, the *Expression module* will combine all the information to choose the adequate linguistic clauses, in order to build a final linguistic report customized to the user's needs.

2.1 Granular Linguistic Model of a Phenomenon (GLMP)

The kernel of the report generator is the *Granular Linguistic Model of a Phenomenon* (GLMP). The designer creates the GLMP as a representation of his/her own perceptions of the monitored phenomenon, organized in several granularity levels.

A GLMP represents our approach to implement CTP [2] [7] [8] to linguistically describe phenomena. This approach represents any type of phenomena, describing its characteristics linguistically. These linguistic variables can be inter-related and the GLMP has to be designed attending to the questions that we want to answer. The linguistic description of a phenomenon evolving in time will have a big number of valid sentences. Each individual sample will represent a picture of the current state of the phenomenon and it will have associated sentences describing it linguistically.

Some of our previous works [3][4][5][6], show detailed definitions of the GLMP. These papers show examples such as the evaluation and descriptions of driving quality in simulators, the description of relevant features of the Mars' surface from satellite images, and the description of the traffic evolution in roads.

The main element of this structure is known as Computational Perception (CP), which is defined by the couple $(A, W) = \{(a_1, w_1), (a_2, w_2), \ldots, (a_n, w_n)\}$. A represents the set of NL sentences that linguistically describes the perception (i.e.: *"Walking speed is* $\{$*slow* $|$ *normal* $|$ *fast*$\}$*".*), and $W \in [0, 1]$ are the validity degrees of each sentence.

On the other hand, the GLMP is a network of Perception Mappings (PM), which are elements used to aggregate and combine CPs. Each PM receives a set of input CPs and transmits upwards an output CP. Each output CP is explained by the PM using a set of input CPs and covers specific aspects of the phenomenon with certain granularity degree. The GLMP corresponding to the practical application can be seen in Fig. 2.

3 Activity Recognition through Mobile Device

This section describes how to apply our approach for linguistic description of complex phenomena to the daily physical activities identification of a person. The relevant modules needed to produce the linguistic report are explained.

3.1 DAQ Module

To identify and report the daily physical activities of a person we have used the triaxial accelerometers embedded in current mobile phones. In particular, we have used *Android* based mobile phones as platform because it is a free and open-source operating system, easy to program. Acceleration data was acquired and stored by an Android application that we have created to run on these devices. Through our application we can control what data is collected as well as how frequently it is done (typically, 50 samples per second). The main idea of this acquisition process is that data are easily obtained and it does not represent an intrusive wearable device for the user. Since mobiles are not always in a fixed position and it is sometimes important to identify in which position the user is, in this paper, we have limited the analysis to the case that the mobile is placed into the trousers pocket. The mobile orientation can be determined using the accelerations module and the azimuth angle.

3.2 Validity Module

In our approach we analyze the accelerations produced in person's daily routine activities, such as walking, standing, sitting, etc. We also approximately calculate the caloric expenditure of each activity carried out. The application output consists of a linguistic summary report with all the relevant features of each daily activity.

Fig. 2 shows the GLMP designed to summarize and highlight the relevant aspects of the physical activities. It is based on first, second and top order CPs that linguistically describe different aspects of the physical activities, and PMs that aggregate and combine them.

The execution of the GLMP and the aggregation rules in its PMs generate the validity degrees of each CP at every instant. Table 1 shows the list of all the CPs used in our application.

Table 1. Table of CPs with the corresponding linguistic variables and labels

CP (y)	**Linguistic Variables**	**Linguistic Labels** (A_y)
$1 - CP_{Phi}$	Azimuth angle variance	{*small, big*}
$1 - CP_{MY'}$	Mean y'	{*low, high*}
$1 - CP_{Var}$	Variance y'	{*low, medium, high*}
$1 - CP_{MZ'}$	Mean z'	{*low, high*}
$2 - CP_S$	Initial State (Detect Azimuth step variations)	{*state 1, state 2*}
$2 - CP_{S1}$	State 1	{*sitting, standing*}
$2 - CP_{S2}$	State 2	{*sitting, traveling, standing, moving while standing*}
$2 - CP_U$	Up state	{*walking, going upstairs, going downstairs*}
$2 - CP_{S2_S}$ & $2 - CP_{U_S}$	Summary CPs	{*short time, some time, rather long time, too much time*} & corresponding labels
$2 - CP_{WS}$	Walking Speed	{*slow, normal, fast*}

At first, some mathematical operations were performed to calculate some relevant magnitudes needed as GLMP inputs. These inputs are represented as Z_1, Z_2, Z_3, Z_4, Z_5 and Z_6:

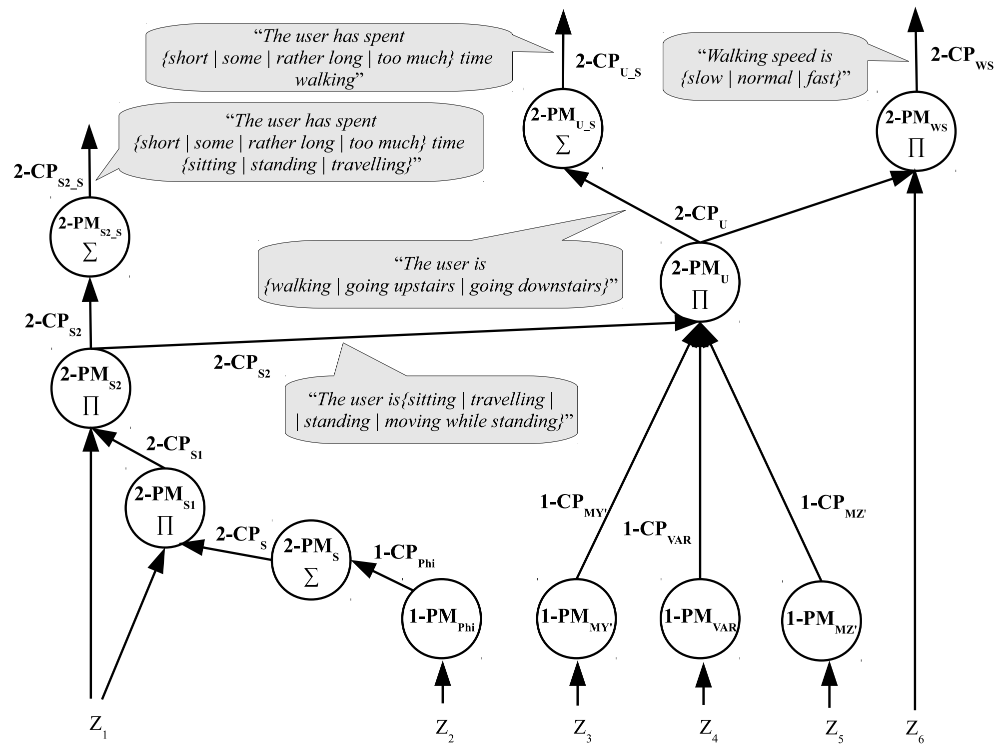


Fig. 2. GLMP for the linguistic description of the physical activities identification

Accelerations Module (*AM*) (Z_6) and Its Variance (Z_1): is calculated using the Pythagorean summing of the three axes values: $AM = \sqrt{x^2 + y^2 + z^2}$.

Azimuth Variance (Z_2): The polar coordinates and the azimuth are used in order to identify relevant changes in mobile orientation and determine mobile's flat and upright positions.

Accelerations *y* (Z_3, Variance Z_4) and *z* (Z_5) in terms of user axes **(*y'* and *z'*):** Once the flat and upright states of the mobiles are distinguished, the accelerations can be transformed into user axes (global coordinate axes).

The validity degrees of the first order perceptions (1-PM) which describe the sensor information ($Z_1, \ldots, Z_6$) are obtained by means of a set of uniformly distributed trapezoidal membership functions, forming strong fuzzy partitions.

The validity degrees of the second order perceptions (2-PM) are calculated by combining or aggregating the information of their input CPs. The symbol Π corresponds to combination functions based on fuzzy IF-THEN rules. The symbol Σ represents the aggregation of 2-PMs, where the validity of the quantified sentences have been computed using the α-cuts method proposed by [9].

3.3 Expression Module

Apart from the goal of obtaining suitable texts to be showed to users, the linguistic reports can be used by therapists with the aim of understanding physical changes, rehabilitation evolutions, patients' habits of life and so on.

The developed application provides two different types of linguistic description reports: a daily report that linguistically and graphically describes the physical activities, their duration and intensity throughout the day, and a periodical report that summarizes physical activity trends throughout a specific period of time (typically one week). In both cases it has been applied basic report templates. See in [5] an example of template that change the structure of the report depending on the validity degrees of the sentences.

The Expression module has the goal of answering the main questions that an expert want to know when he/she is going to design and elaborate a treatment program or giving medical assessment. The GLMP has to be capable to provide to the *Expression module* the information that it needs to answer these questions.

4 Experimentation

In order to validate the accuracy of the developed system we believed advisable to perform a series of test by contrasting the real activity sequences with the application output ones.

The first experiment consisted of developing 10 tests of approximately one hour. For each test, one user recorded the accelerations produced by his/her mobile phone and took note of the activities performed during the session (normal or routine activities were carried out, no specific ones). The matching between the performed and reported activities allows us to know the accuracy of the application in this task.

Table 2 shows the results of the experiment. First columns shows the name and times of the recorded sessions. Second columns counts the number of activities performed in each session. Columns 3 and 4 present the number of correctly and incorrectly classified activities respectively. Last column shows the percentage of time the activity identification is correct during the test. The percentage error calculation penalizes not only the fact of misclassification of an activity but also the amount of time this classification is wrong.

Table 2. Validation experiments and classification results

Test names & duration	Num. of activities	Correctly classified activities	Misclassified activities	% of time correctly classified
T-1 (58')	10	10	1 (1')	98.28
T-2 (96')	9	9	0	100
T-3 (45')	7	7	1 (1')	97.77
T-4 (65')	8	8	0	100
T-5 (58')	11	10	1 (1')	98.28
T-6 (40')	11	11	0	100
T-7 (63')	8	8	0	100
T-8 (48')	5	5	0	100
T-9 (38')	7	7	0	100
T-10 (63')	9	8	1 (1')	98.41

The average of correct classifications was equal to **99.27%**, which means a very high accuracy. Incorrectly classified activities typically corresponds to minor mistakes, mostly produced by strange accelerations or movements existing in the

Table 3. Example of validation experiment. Activities performed vs. Activities reported.

List of Performed Activities	Linguistic Report of Activities
"Start application with mobile phone on a table"	*"19:00 - 19:18 Out (✓)"*
"Introduce mobile into the pocket and stand for a few minutes"	*"19:18 - 19:19 Standing (✓)"*
"Walk some minutes at home"	*"19:19 - 19:21 Walking (✓)"*
"Go downstairs to the garage"	*"19:21 - 19:21 Down stairs (✓)"*
"Walk to the car"	*"19:21 - 19:22 Walking (✓)"*
"Stand beside the car keeping some things"	*"19:22 - 19:23 Standing (✓)"*
"Get in the car and start driving"	*"19:23 - 19:38 Traveling (✓)"*
"Arrive destination, leave car and walk some minutes"	*"19:38 - 19:41 Walking (✓)"*
"Walk upstairs several floors"	*"19:41 - 19:42 Climbing stairs (✓)"*
"Stand some minutes"	*"19:43 - 19:43 Standing (✓)"*
"Sit down, and stop application after some minutes"	*"19:43 - 19:49 Sitting (✓)"*

course of normal life. These minor mistakes are not relevant when producing a linguistic report of conclusions.

Table 3 shows an example of one test (T-6 in Table 2) used for the validation. It presents the list of performed activities annotated by the user and the corresponding linguistic report produced by the application. The state *'Out'* corresponds to the state when the mobile phone is not with the user.

The main feature of this work is that the application has the ability of providing a wide variety of reports, highlighting different characteristics of human activity. It will depend on the particular needs of the different users. For example, we could customize the output reports considering the amounts of time the user should be walking or sitting, as well as the quantity of energy he/she must spend. This configuration set could be done by an expert in order to advice patients on behalf of their activity levels. The generated reports could be similar to the following one:

"During these 40 minutes you walked for a short time and you were sitting most of the time. You have burned very few calories. In order to meet your objectives, you should increase your activity level."

The second experiment consists in obtaining the accelerations generated for a whole day over the whole week. Each daily recording period has consisted of 10 hours of continuous recording. The main goal of this experimentation resided in demonstrating that significant conclusions can be drawn from the observation of long periods of time by comparing results, trends and so on. The report obtained from the observation of 7 experimental days was the following:

"During this week you have consumed less energy than it was expected. The energy consumption has been very similar every day of the week."
"The day of the week you have walked less time was on Thursday, however, on Sunday you have been walking far longer than the other days."
"On Tuesday and Saturday, you have spent much more time standing than other days. In addition, you spend too much time sitting, so you should reduce as much as possible the time spent sitting."

5 Concluding Remarks

In this paper we have focused our efforts in the development of a GLMP which describes the most common physical activities of a person during a day. We have developed a practical application that lies in the use of current mobile phones to acquire and store user accelerations in an inexpensive and non-intrusive way. Experimentation allows us to demonstrate the viability of our approach to become a real commercial application.

In this paper we present a glimpse of our ongoing work devoted to generate linguistic descriptions of human activities, being a field constantly growing and with a great potential to be improved. Detecting falls, combining acceleration data with other sensors information, or incorporating the capability of interacting in social networks would be some of the improvements that we want to tackle in the future.

Acknowledgment. This work has been funded by the Spanish Government (MICINN) under project TIN2011-29827-C02-01.

References

1. Zadeh, L.A.: The concept of linguistic variable and its application to approximate reasoning. Information sciences 8, 199–249 (1975)
2. Zadeh, L.A.: From computing with numbers to computing with words - from manipulation of measurements to manipulation of perceptions. IEEE Transactions on Circuits and Systems 45 (1999)
3. Eciolaza, L., Trivino, G., Delgado, B., Rojas, J., Sevillano, M.: Fuzzy linguistic reporting in driving simulators. In: Proceedings of the 2011 IEEE Symposium on Computational Intelligence in Vehicles and Transportation Systems, Paris, France, pp. 30–37 (2011)
4. Eciolaza, L., Trivino, G.: Linguistic Reporting of Driver Behavior: Summary and Event Description. In: 11th International Conference on Intelligent Systems Design and Applications (ISDA), Córdoba, Spain, pp. 148–153 (2011)
5. Alvarez-Alvarez, A., Sanchez-Valdes, D., Trivino, G.: Automatic Linguistic Description about Relevant Features of the Mars' Surface. In: Proceedings of the 11th International Conference on Intelligent Systems Design and Applications (ISDA), Córdoba, Spain, pp. 154–159 (2011)
6. Alvarez-Alvarez, A., Sanchez-Valdes, D., Trivino, G., Sánchez, A., Suárez, P.D.: Automatic linguistic report about the traffic evolution in roads. Expert Systems with Applications 39, 11293–11302 (2012)
7. Zadeh, L.A.: A new direction in AI. Towards a Computational Theory of Perceptions of measurements to manipulation of perceptions. AI Magazine 22 (2001)
8. Zadeh, L.A.: Toward Human Level Machine Intelligence - Is It Achievable? The Need for a Paradigm Shift IEEE Computational Intelligence Magazine (2008)
9. Delgado, M., Sánchez, D., Vila, M.: Fuzzy cardinality based evaluation of quantified sentences. Int. Journal of Approximate Reasoning 23, 23–66 (2000)

7.2. Increasing the Granularity Degree in Linguistic Descriptions of Quasi-periodic Phenomena

D. Sanchez-Valdes, and G. Trivino. "Increasing the Granularity Degree in Linguistic Descriptions of Quasi-periodic Phenomena". *Flexible Query Answering Systems*, pp. 281–292, Springer.

Increasing the Granularity Degree in Linguistic Descriptions of Quasi-periodic Phenomena

Daniel Sanchez-Valdes and Gracian Triviño

European Centre for Soft Computing (ECSC)
Mieres, Asturias, Spain
{daniel.sanchezv,gracian.trivino}@softcomputing.es

Abstract. In previous works, we have developed some computational models of quasi-periodic phenomena based on Fuzzy Finite State Machines. Here, we extend this work to allow designers to obtain detailed linguistic descriptions of relevant amplitude and temporal changes. We include several examples that will help to understand and use this new resource for linguistic description of complex phenomena.

Keywords: Linguistic description of data, Computing with Perceptions, Fuzzy Finite State Machine, Quasi-periodic phenomena.

1 Introduction

Computational systems allow obtaining and storing huge amounts of data about phenomena in our environment. Currently, there is a strong demand for computational systems that can interpret and linguistically describe the large amount of information that is being generated.

Some physical phenomena provide signals with a similar repetitive temporal pattern. These phenomena are called quasi-periodic due to their variations in period and amplitude. Examples of this type of signals are electrocardiograms, the breathing, accelerations produced during the human gait, vibrations of musical instruments, etc. Popular approaches to deal with quasi-periodic phenomena vary from Wavelets transform [1] to Hidden Markov Models [2] and Neural Networks [3]. Fuzzy Finite State Machines (FFSM) are specially useful tools to model dynamical processes that change in time, becoming an extension of classical Finite State Machines [4][5].

Our research line is based on the Computational Theory of Perceptions introduced by Zadeh [6][7]. In previous works, we have developed computational systems able to generate linguistic descriptions of different types of phenomena, e.g., gait analysis [8], activity recognition [9] and traffic evolution [10][11].

We have used FFSMs to model and linguistically describe the temporal evolution of quasi-periodic signals during a period of time. In this work, our goal consists of exploring the possibility of generating linguistic descriptions of those instants in which phenomena are significantly deviated from the model. Here, we identify each state and linguistically report the evolution of the phenomenon.

H.L. Larsen et al. (Eds.): FQAS 2013, LNAI 8132, pp. 281–292, 2013.

When the input signal does not completely suit with the available model, the computational system describes the reasons that cause this event, providing a finer monitoring of the input signal. This analysis is appropriate to overcome signal processing, predictive control and monitoring tasks.

This paper is organized as follows: Section 2 describes the architecture of computational systems able to create linguistic descriptions of complex phenomena. Section 3 explains how to apply it to the study of quasi-periodic signals. Section 4 shows a set of simple examples that demonstrate the potential of our approach. Finally, Section 5 provides some concluding remarks and future works.

2 Linguistic Description of Complex Phenomena

This section briefly describes the main concepts of our approach to linguistic description of complex phenomena.

2.1 Computational Perception (CP)

A CP is a tuple (A, W) described as follows:

$A = (a_0, a_1, a_2, \ldots, a_n)$ is a vector of linguistic expressions (words or sentences in Natural Language) that represents the whole linguistic domain of CP. The components of A are defined by the designer by extracting the most suitable sentences from the typically used ones in the application domain of language. In the application context, each component a_i describes the linguistic value of CP in each situation of the phenomenon with specific granularity degree. These sentences can be either simple, e.g., $a_i =$ *"The temperature is quite high"* or more complex, $a_i =$ *"Today the weather is better than the last days"*. If the perception does not match with any of the available possibilities the model uses the linguistic expression a_0, e.g, *"The available model cannot explain completely the current situation"*.

$W = (w_0, w_1, w_2, \ldots, w_n)$ is a vector of validity degrees $w_i \in [0, 1]$ assigned to each a_i. In the application context, w_i represents the suitability of a_i to describe the perception. After including a_0, the components of A form a strong fuzzy partition of the domain of existence of CP, i.e., $\sum_{i=0}^{n} w_i = 1$.

2.2 Perception Mapping (PM)

We use PMs to combine or aggregate CPs. A PM is a tuple (U, y, g, T) where:

U is a vector of input CPs, $U = (u_1, u_2, \ldots, u_m)$, where u_i are tuples (A_i, W_i) and m the number of input CPs. We call *first order perception mappings* (1PMs) when U are numerical values obtained , e.g., from a database.

y is the output CP, $y = (A_y, W_y)$.

g is an aggregation function $W_y = g(W_1, W_2, \ldots, W_m)$, where W_i are the vectors of validity degrees of the m input CPs. In Fuzzy Logic, many different types of aggregation functions have been developed. For example, g can be implemented using a set of fuzzy rules.

T is a text generation algorithm that allows generating the sentences in A_y. In our current approach, T is typically a basic linguistic template, e.g., *"Likely the temperature sensor is wrong"* and/or *"The temperature of this room is* $\{$*high* | *medium* | *low*$\}$*"*.

2.3 Granular Linguistic Model of Phenomena (GLMP)

GLMP consists of networks of PMs (see Fig. 2). We say that output CPs are explained by PMs using a set of input CPs. In the network, each CP covers an specific aspect of the phenomenon with certain granularity degree.

We call *first order computational perceptions* (1CPs) to those ones obtained from the system input and we call *second order computational perception* (2CPs) to those ones explained by previous CPs. By means of using different aggregation functions and different linguistic expressions, the GLMP paradigm allows the designer to model computationally her/his perceptions. Note that, after being instantiated with a set of input data, the GLMP provides a structure that, in medium size applications, could include hundreds of valid sentences. We will see in the following sections how to merge these sentences in a template to generate linguistic reports.

3 Linguistic Modeling of Quasi-periodic Phenomena

In order to illustrate the method for creating linguistic descriptions of quasi-periodic phenomena, we use a simple example built around a sinusoidal signal $u(t) = sin(\omega t)$. Fig. 1 shows the modeled signal $u(t)$, which is divided into four states q_1, q_2, q_3 and q_4. Each state q_i is repeated along the time and has a set of special characteristics that distinguish it from the other states.

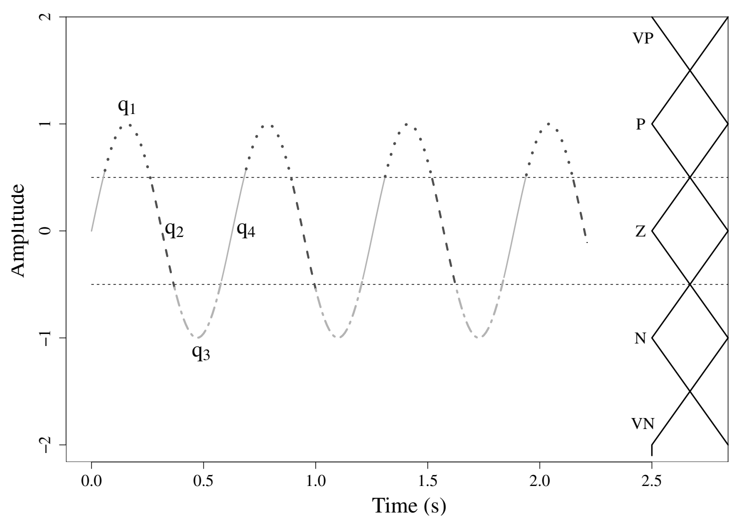


Fig. 1. Sinusoidal signal with related fuzzy labels

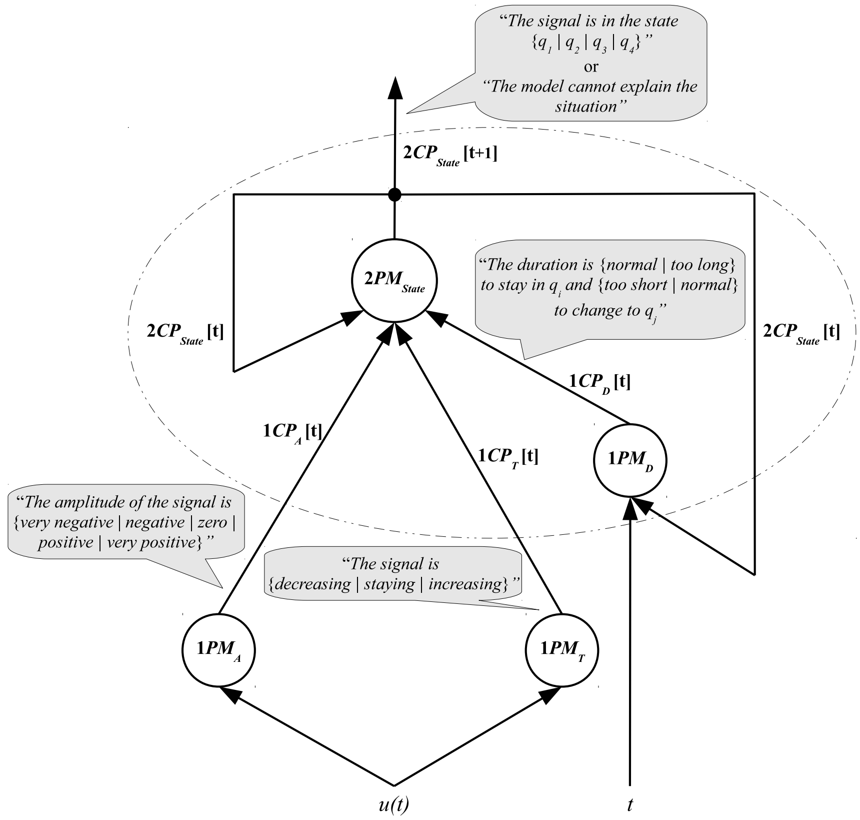


Fig. 2. GLMP that models the signal

Fig. 2 shows a GLMP designed to model this signal. $1CP_A$ and $1CP_T$ describe the amplitude and trend of the input signal respectively. $2CP_{FFSM}$ shows the state of the signal at every time instant t and $1CP_D$ presents the duration of the signal in each state q_i.

3.1 Signal Amplitude ($1PM_A$)

U are the numerical values ($u[t] \in \mathbb{R}$) of the input signal $u[t] = sin(\omega t)$, where $\omega = 10$ and t takes values in the time interval $t \in [0, 8\pi]$. Therefore, the period of the signal is $K = 2\pi/10$.

y is the output $1CP_A$, which describes the possible values of the signal amplitude with the vector A =(*Very negative (VN), Negative (N), Zero (Z), Positive (P), Very positive (VP)*).

g is the function that calculates the validity degrees of the output CP. These values are obtained by means of uniformly distributed triangular membership functions (MFs) forming strong fuzzy partitions. Here, the five linguistic labels associated to each expression are represented with triangular membership functions defined by their vertexes as follows: $\{VN\ (-\infty, -2, -1)$, N $(-2, -1, 0)$, Z $(-1, 0, 1)$, P $(0, 1, 2)$, VP $(1, 2, \infty)\}$. This linguistic labels are represented in Fig. 1, together with the input signal $u(t)$.

T is the text generator that produces linguistic expressions as follows: *"The amplitude of the signal is {very negative | negative | zero | positive | very positive}"*.

3.2 Signal Trend ($1PM_T$)

U are, as in $1PM_A$, the numerical values ($u[t] \in \mathbb{R}$) of the input signal.
y is the output $1CP_T$, which describes the possible values of the signal trend with the vector A =(*Decreasing (D), Staying (S), Increasing (I)*).
g is the output function that calculates the signal trend as $u[t] - u[t-1]$. It is defined by the vertexes of triangular MFs as follows: $\{D\ (-\infty, -0.01, 0),\ S\ (-0.01, 0, 0.01),\ I\ (0, 0.01, \infty)\}$.
T produces linguistic expressions defined as follows: *"The signal is {decreasing | staying | increasing}"*.

3.3 State Duration ($1PM_D$)

U are the input $2CP_{State}$ and the time instant t. Each state q_i has an associated duration d_i, which is numerically calculated with the time t.
y is the output $1CP_D$ which describes how long the phenomenon is in each state q_i. The possible linguistic expressions to describe this perception are contained in the vector: A =(*too short to change to* q_j, *normal to change to* q_j, *normal to stay in* q_i, *too long to stay in* q_i).
g is the output function that calculates the validity degrees of $1CP_D$. When the state q_i takes the value zero, its duration d_i takes value zero. However, when q_i takes values bigger than zero, its duration d_i increase its value according to the time, measuring the duration of the state.

Based on the duration d_i of each state we define two temporal conditions described as follows:

The **time to stay** is the maximum time that the signal is allowed to stay in state i. We associate two linguistic labels that describe this perception: {*normal to stay in* q_i, *too long to stay in* q_i}.

The **time to change** is the minimum time that the signal must be in state i before changing to state j. We associate other two linguistic labels that describe this perception: {*too short to change to* q_j, *normal to change to* q_j}.

Fig. 3 shows the linguistic labels used to define these temporal constraints. In this application, states q_1 and q_3 have the same duration and the linguistic labels are trapezoidal MFs that can be defined by their vertexes as follows:
Time to stay: {*normal* $(-\infty, -\infty, 5/12, 1/2)$, *too long* $(5/12, 1/2, \infty, \infty)$}.
Time to change: {*too short* $(-\infty, -\infty, 1/4, 5/12)$, *normal* $(1/4, 5/12, \infty, \infty)$}.
On the other hand, states q_2 and q_4 share the same duration and the linguistic labels can be defined by their vertexes as follows:
Time to stay: {*normal* $(-\infty, -\infty, 1/3, 1/2)$, *too long* $(1/3, 1/2, \infty, \infty)$}.
Time to change: {*too short* $(-\infty, -\infty, 1/6, 1/4)$, *normal* $(1/6, 1/4, \infty, \infty)$}.

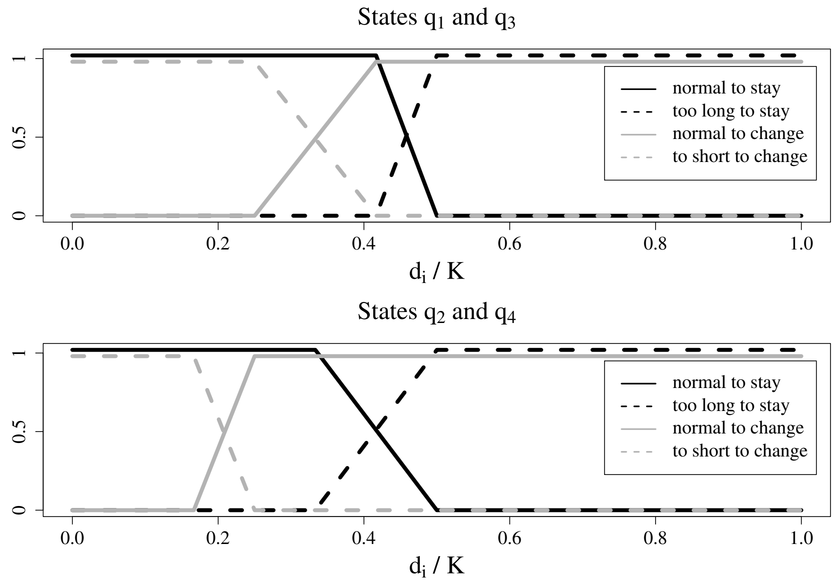


Fig. 3. Times to stay and times to change of states q_1, q_2, q_3 and q_4

These vertexes are expressed in terms of K as a percentage of the signal period. They have been carefully calculated attending to the signal structure. For example, if the signal is in the state q_1 and, at specific time t, d_1 is equal to 0.2 ($d_1/K = 0.318$), the aggregation function g indicates that this duration is *normal to stay in* q_1 (1), *too long to stay in* q_1 (0), *normal to change to* q_2 (0.4) and *too short to change to* q_2 (0.6).

T produces linguistic expressions as follows: *"The duration is* {*normal, too long*} *to stay in* q_i *and* {*too short, normal*} *to change to* q_j*"*, with $i \in \{1, 2, 3, 4\}$.

3.4 Signal State ($\boldsymbol{2PM_{State}}$)

With the information about the amplitude, trend, duration of each state and the signal state in t, this $2PM_{State}$ calculates the next state of the signal in $t + 1$.

U are the input CPs: $\{1CP_A$, $1CP_T$, $1CP_D$, $2CP_{State}\}$

y is the output $2CP_{State}$, that describes the possible states of the signal as the states of a FFSM. The vector of linguistic expressions is $A = (q_0, q_1, q_2, q_3, q_4)$, where q_0 means that the input signal does not fit with any of the other fuzzy states (see Fig. 4). The definition of the number of states and allowed transitions have to be designed according to the modeled signal, i.e., the higher the complexity of the monitored signal the bigger the number of states to monitor its evolution fine.

g is the aggregation function implemented using a set of fuzzy rules. We have used twelve fuzzy rules, namely, R_{ii} to remain in the state i and R_{ij} to change from state i to state j. The whole rules base is listed as follows:

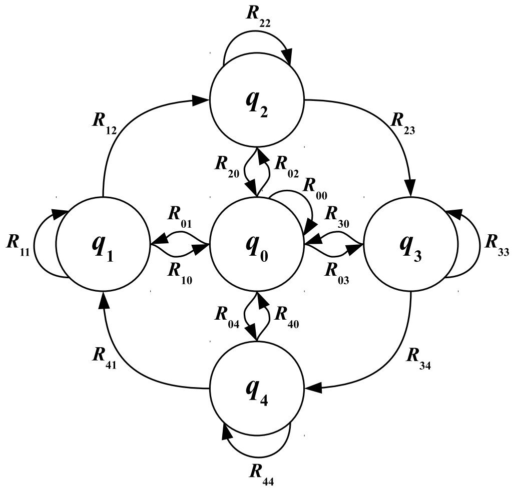


Fig. 4. State diagram of the FFSM for the sinusoidal signal

R_{11}: IF ($State[t]$ is q_1) AND (*amplitude* is P) AND (*duration* is *normal to stay in* q_1) THEN ($State[t+1]$ is q_1)
R_{22}: IF ($State[t]$ is q_2) AND (*amplitude* is Z) AND (*trend* is D) AND (*duration* is *normal to stay in* q_2) THEN ($State[t+1]$ is q_2)
R_{33}: IF ($State[t]$ is q_3) AND (*amplitude* is N) AND (*duration* is *normal to stay in* q_3) THEN ($State[t+1]$ is q_3)
R_{44}: IF ($State[t]$ is q_4) AND (*amplitude* is Z) AND (*trend* is I) AND (*duration* is *normal to stay in* q_4) THEN ($State[t+1]$ is q_4)
R_{12}: IF ($State[t]$ is q_1) AND (*amplitude* is Z) AND (*trend* is D) AND (*duration* is *normal to change to* q_2) THEN ($State[t+1]$ is q_2)
R_{23}: IF ($State[t]$ is q_2) AND (*amplitude* is N) AND (*trend* is D) AND (*duration* is *normal to change to* q_3) THEN ($State[t+1]$ is q_3)
R_{34}: IF ($State[t]$ is q_3) AND (*amplitude* is Z) AND (*trend* is I) AND (*duration* is *normal to change to* q_4) THEN ($State[t+1]$ is q_4)
R_{41}: IF ($State[t]$ is q_4) AND (*amplitude* is P) AND (*trend* is I) AND (*duration* is *normal to change to* q_1) THEN ($State[t+1]$ is q_1)
R_{01}: IF ($State[t]$ is q_0) AND (*amplitude* is P) AND (*trend* is I) THEN ($State[t+1]$ is q_2)
R_{02}: IF ($State[t]$ is q_0) AND (*amplitude* is Z) AND (*trend* is D) THEN ($State[t+1]$ is q_3)
R_{03}: IF ($State[t]$ is q_0) AND (*amplitude* is N) AND (*trend* is D) THEN ($State[t+1]$ is q_4)
R_{04}: IF ($State[t]$ is q_0) AND (*amplitude* is Z) AND (*trend* is I) THEN ($State[t+1]$ is q_1)
R_{i0}: ELSE ($State[t+1]$ is q_0)

T produces linguistic expressions that be adapted depending on the situation.
When the signal is in the states q_1, q_2, q_3 or q_4 the template is the following:

"The signal is in the state $\{q_1,\ q_2,\ q_3,\ q_4\}$*"*. However, when the signal is in the state q_0 the system reports the following expression: *"The model cannot explain the current situation"*.

3.5 Remarks

Following the works developed in [8][11], in this paper we present in detail the CPs existing into the FFSM. $2PM_{States}$ calculates the signal state at each instant t. The output $2CP_{States}$ deals with several goals: first, it allows to generate linguistic descriptions about each state q_i of the input signal; second, it feeds the $2PM_{States}$ to calculate the next state taking into account the current state at t; and finally, it feeds the $1PM_D$ to calculate the temporal constraints of the FFSM.

It is worth remarking that this model represents the general form of a time-invariant discrete system in the state space, formulated as:

$$\begin{cases} X[t+1] = f(X[t], U[t]) \\ Y[t] = g(X[t], U[t]) \end{cases}$$

where:

- U is the input vector of the system: $(u_1, u_2, ..., u_{n_u})$, with n_u being the number of input variables.
- X is the state vector: $(x_1, x_2, ..., x_n)$, with n being the number of states.
- Y is the output vector: $(y_1, y_2, ..., y_{n_y})$, with n_y being the number of output variables.
- f is the function which calculates the state vector at time step $t+1$.
- g is the function which calculates the output vector at time step t.

In our approach, the model could be formulated as:

$$2CP_{State}[t+1] = f(2CP_{State}[t], U[t])$$

where, $2CP_{State}$ is the state vector at time step t, U is the input vector composed by the input variables $1CP_A$, $1CP_T$ and $1CP_D$, and f is the function that calculates the state vector at time step $t+1$. In this application, we have used the state vector as output vector and we have not used g.

4 Examples

In order to illustrate how the model works, Figs. 6, 7, 8 and 9 show different behaviors of the text generator depending on the input signal values. Each figure represents the input signal. Below of the signal is illustrated the evolution of each state q_i (w_{q_i}). The vertical line indicates the time instant where the description is generated. As mentioned before, the modeled FFSM was designed to recognize the states of an input signal of the form $u(t) = sin(10t)$.

The main objective is to linguistically describe the evolution of the signal, i.e., to indicate the state q_i in which the signal is at every time instant and, in case of deviation from model, to specify the details. Note that the way to analyze the reasons why the model cannot explain the signal in a specific period consists of going over the GLMP in the reverse direction the model was built.

The linguistic reports obtained can be used by experts to understand changes in the signal and foreseeing its future behavior. We have applied the report template shown in Fig. 5. This template changes the report depending on the validity degrees of the sentences. The sentences with highest validity degree are chosen at each time instant for each CP.

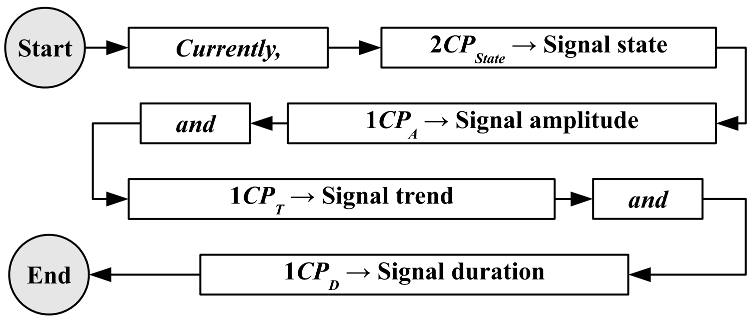


Fig. 5. Template for the linguistic report

4.1 Example 1

Fig. 6 shows the fuzzy states evolution when the input signal is exactly the expected one, i.e., when the set of rules represent perfectly its evolution. Therefore, w_{q_0} will be equal to zero along the time. A example of instant report when the signal matches completely the model is as follows

"Currently, the signal is in the state q_1. The amplitude of the signal is positive, it is increasing, and the duration is normal to stay in q_1 and too short to change to q_2".

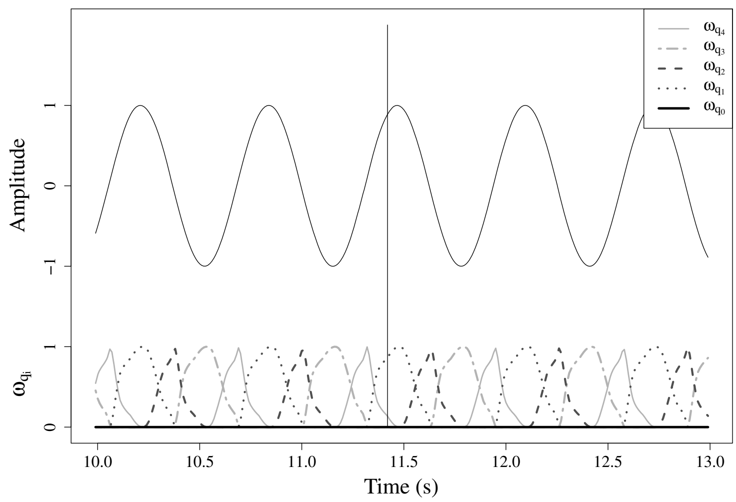


Fig. 6. States when the signal is exactly the expected one

4.2 Example 2

When the input signal has a perturbation or an unforeseen situation, the set of rules does not recognize its behavior and the model indicates that the signal is in an uninterpretable state (q_0). The activation of the state q_0 informs that the input signal does not match with the typical sinusoidal signal. Fig. 7 shows a first example of this type of situations. In this case, the input signal has a perturbation that maintains its value constant during a period of time. At the time of failure, the signal was in the state q_1 and the temporal constraint related to the maximum time that the signal is allowed to be in this state force the system to finish it. As the conditions to be in next state were not fulfilled, the system indicates that the signal is in the state q_0. The linguistic description obtained when the perturbation occurs is as follows:

*"Currently, the model cannot explain the situation. The amplitude of the signal is positive, it is **staying**, and the duration is **too long** to stay in q_1 and normal to change to q_2".*

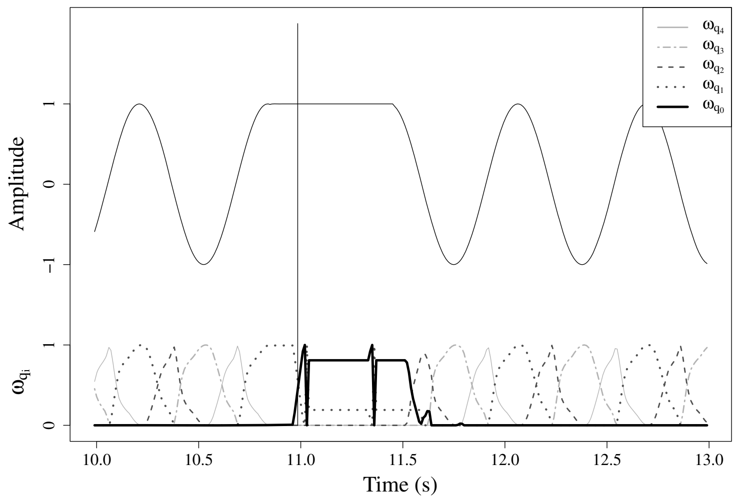


Fig. 7. States when the signal has a temporal error

4.3 Example 3

In this example, another perturbation modifies the normal behavior of the input data (Fig. 8). In this case, the state q_1 decreases its duration, being in this state less than a half its normal duration. Again, the temporal conditions of the set of rules detect that the input signal does not fit well with the model in this instant. The reason is different with respect to the previous example and now the linguistic description obtained is as follows:

*"Currently, the model cannot explain the situation. The amplitude of the signal is zero, it is decreasing, and the duration is normal to stay in q_1 and **too short** to change to q_2".*

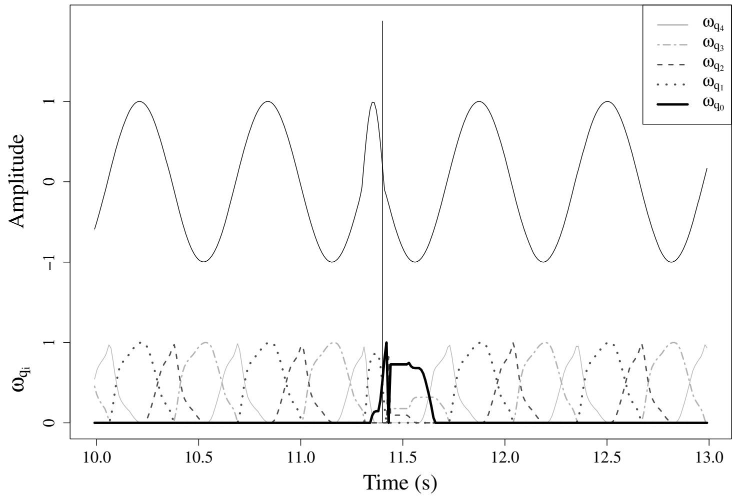


Fig. 8. States when the signal has another type of temporal error

4.4 Example 4

This last example shows the modeling of a different type of perturbation (Fig. 9). In this case the error is produced by an amplitude perturbation that forces the input signal to be almost twice bigger than the expected one during the state q_3. The conditions that model the signal amplitude detect a strange behavior. This is reflected in the evolution of q_0, which takes high values in these moments. The linguistic description obtained is as follows:

"Currently, the model cannot explain the situation. The amplitude of the signal is ***very negative****, it is decreasing, and the duration is normal to stay in q_3 and too short to change to q_1"*

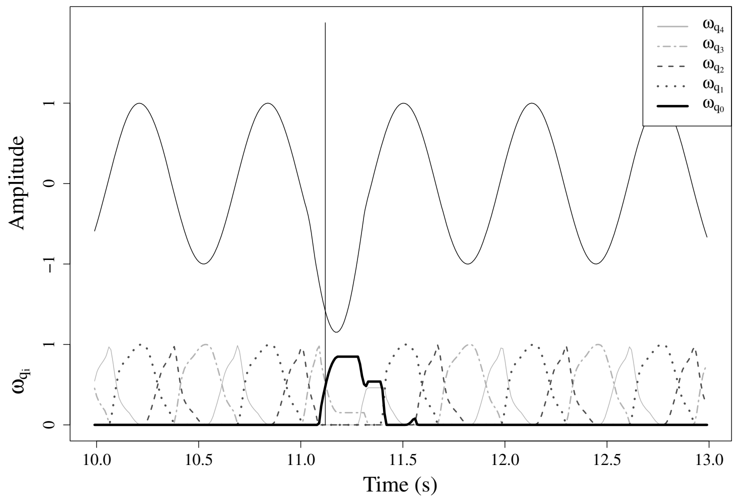


Fig. 9. States when the signal has an amplitude error

5 Conclusions

We have developed previous research works by exploring the possibility of providing detailed linguistic descriptions of quasi-periodic phenomena. Here, we focus on the description of relevant deviations of the signal from the available computational model.

There are several possibilities for pending work, e.g., the use of quantifiers to summarize the frequency of happening of these perturbations.

An important number of applications could take advantage of the contribution of this research line. For example, the results that we present here can be directly applied to analyze and describe anomalies in physiological signals, such as electrocardiogram and human gait.

Acknowledgment. This work has been funded by the Spanish Government (MICINN) under project TIN2011-29827-C02-01.

References

1. Burrus, C.S., Gopinath, R.A., Guo, H.: Introduction to wavelets and wavelet transforms: a primer. Prentice Hall, pper Saddle River (1998)
2. Rabiner, L.R.: A tutorial on hidden Markov models and selected applications in speech recognition. Proceedings of the IEEE 77(2), 257–286 (1989)
3. Ripley, B.D.: Pattern recognition and neural networks. Cambridge university press (2008)
4. Lawson, M.V.: Finite Automata. Chapman and Hall/CRC (2003)
5. Davis, J.: Finite State Machine Datapath Design, Optimization, and Implementation (Synthesis Lectures on Digital Circuits and Systems). Morgan and Claypool Publishers (2008)
6. Zadeh, L.A.: From computing with numbers to computing with words - from manipulation of measurements to manipulation of perceptions. IEEE Transactions on Circuits and Systems - I: Fundamental Theory and Applications 45(1), 105–119 (1999)
7. Zadeh, L.A.: A new direction in AI. Towards a Computational Theory of Perceptions of measurements to manipulation of perceptions. AI Magazine 22 (2001)
8. Alvarez-Alvarez, A., Trivino, G.: Linguistic description of the human gait quality. Engineering Applications of Artificial Intelligence (2012)
9. Sanchez-Valdes, D., Eciolaza, L., Trivino, G.: Linguistic Description of Human Activity Based on Mobile Phone's Accelerometers. In: Bravo, J., Hervás, R., Rodríguez, M. (eds.) IWAAL 2012. LNCS, vol. 7657, pp. 346–353. Springer, Heidelberg (2012)
10. Eciolaza, L., Trivino, G., Delgado, B., Rojas, J., Sevillano, M.: Fuzzy linguistic reporting in driving simulators. In: Proceedings of the 2011 IEEE Symposium on Computational Intelligence in Vehicles and Transportation Systems, Paris, France, pp. 30–37 (2011)
11. Alvarez-Alvarez, A., Sanchez-Valdes, D., Trivino, G., Sánchez, Á., Suárez, P.D.: Automatic linguistic report about the traffic evolution in roads. Expert Systems with Applications 39(12), 11293–11302 (2012)

7.3. Dynamic linguistic descriptions of time series applied to self-track the physical activity

D. Sanchez-Valdes, A. Alvarez-Alvarez, and G. Trivino. "Dynamic linguistic descriptions of time series applied to self-track the physical activity". Submitted in *Fuzzy Sets and Systems*. Accepted with minor revisions. May 2015.

Dynamic linguistic descriptions of time series applied to self-track the physical activity

Daniel Sanchez-Valdes[1,*], Alberto Alvarez-Alvarez[1], Gracian Trivino[1]

Abstract

Self-tracking our physical activity allows us to acquire a better self-knowledge about our physical, and even mental health condition. Currently, the description of the time series provided by sensors is done by means of graphics and tables, that are hard to interpret by non-expert humans. Here, we present a computational application that dynamically describes in natural language the physical activity. The final reports are adapted to the everyday language and user's needs. The application highlights the relevant information obtained at different levels of temporal detail. We have included experimental results that demonstrate the flexibility and applicability of the new tool.

Keywords: Dynamic report, Self-tracking, Linguistic description of time series

1. Introduction

Current smartphones, among other electronic devices, contain diverse and powerful sensors, e.g., GPS, microphones, cameras, compasses, accelerometers and gyroscopes that generate big amounts of time series data. These data are related to the user's physical activities and habits, opening up a wide and exciting field of research in self-tracking or life-logging areas.

The topic of self-tracking has emerged in recent years thanks to the technology advances. It can be successfully applied to address health concerns that arise due to inactivity, e.g., cardiovascular diseases, hypertension, osteoporosis, and the critical public health threat of childhood obesity [1]. It allows users to acquire a better self-knowledge, as well as it allows experts (either endocrinologists, psychologists or personal trainers, among others) to adjust treatments and analyze the physical evolution of their patients, e.g., when analyzing results of medical interventions.

Some activity recognition systems have been developed using multiple accelerometers strapped to the subject's extremities [2]. A good number of activities are recognizable but it is impractical to continuously wear so many accelerometers over the body in daily life. In [3], authors used only a waist mounted accelerometer to estimate the energy expenditure. Authors in [4] exploit the capabilities of current smartphones to use the accelerations taken from these devices situated in the user's trousers pocket, recognizing daily activities such as walking, running, climbing stairs, sitting and stopped. In [5], a living activity recognition is performed by wearing the mobile device in the breast pocket, in order to determine whether the user was walking, quiet or performing a task. The study of these solutions shows the need of developing algorithms that do not require a specific location of the device.

Most smartphone based activity recognition systems apply statistical machine learning models to identify and label the activities, like Bayesian Networks [6], Gaussian Mixture Models [7], Hidden Markov Models

*Corresponding author

Email addresses: `daniel.sanchezv@softcomputing.es` (Daniel Sanchez-Valdes), `alberto.alvarez@softcomputing.es` (Alberto Alvarez-Alvarez), `gracian.trivino@softcomputing.es` (Gracian Trivino)

[1]European Centre for Soft Computing, c/ Gonzalo Gutierrez Quirós, s/n, 33600 Mieres, Asturias, Spain (+34 985 45 65 45).

Preprint submitted to Fuzzy Sets and Systems *May 7, 2015*

[8, 9], Dynamic Bayesian Networks [10], Support Vector Machines [11, 12], Artificial Neural Networks [13] or Decision Trees [14]. Most of these machine learning models infer results without providing any knowledge or explanation of they internal working (black box systems). The development of a system with a "transparent" internal working could be really interesting from the interpretability point of view.

Apart from data collection and interpretation, the self-tracking process also consists of communication. In order to be useful, the time series obtained from the smartphone sensors must be represented in an understandable way, by giving in each type of situation their relationship to data context and, in general, the information related to each specific physical activity.

Usually, the descriptions of this type of data are based on reports that contain text and graphics produced by human experts. There are also computational systems that automatically analyze data and provide reports based on statistical information, i.e., numerical data accompanied by graphics and tables. In this sense, nowadays there is an emerging offer of commercial products [15, 16] that monitor the physical activity, whose goal is to measure some physical parameters (steps, calories, distance, sleep quality, heart rate and so on) but the generated reports are far from being interpretable and high quality reports.

According to the Grice's Maxims [17], a good quality report must meet the four maxims or principles listed as follows:

1) It has to be truth, not saying that which lacks adequate evidence (Maxim of Quality).

2) It has to be clear, avoiding ambiguity and being orderly (Maxim of Manner).

3) It has to be relevant (Maxim of Relation).

4) And, finally, it has to have an adequate extension, not being more informative than is required (Maxim of Quantity).

There is a strong demand for computational systems that can interpret and linguistically describe the information, adapting the relevant linguistic reports according to the user's needs. The topic of linguistically summarize time series has been tackled by many authors. One possible approach to describe changes in the behavior of time series is by using fuzzy logic [18], thanks to its ability for linguistic concept modeling and semantic expressiveness. Authors in [19] used fuzzy rules to perform linguistic descriptions of time series by using several parameters such as the degree of truth, coverage and reliability. There are several approaches to explain the frequency of trends that measure the duration and variability of changes within a time series by means of linguistic summaries [20]. They have been successfully applied to summarize long term trends about one dimensional data about motion and restlessness from an elder person during 15 months [21]. Finally, there are some recent works like [22], where authors introduced a new approach to linguistic summarization of time series based on the use of a fuzzy hierarchical partition of the time dimension and the evaluation of quantified sentences. In previous works, we have generated assessing reports in truck driving simulators [23], reports about traffic evolution in roads [24], about the relevant features of the Mars' surface [25] and linguistic descriptions of visual double stars [26]. In addition, we have worked with accelerometer data by automatically generating linguistic reports about human gait quality [27], gesture recognition [28] and activity recognition [29, 30]. In this paper, we merge the results of these research works with our own work in the field of linguistic description of complex phenomena to design a practical tool that provides relevant linguistic descriptions of time series in order to self-track the physical activity, according to the Grice's Maxims.

This paper is organized as follows. First, Section 2 presents the main components of the extended architecture. Section 3 describes the designed model that provides valid and relevant sentences to self-track the physical activity. Section 4 explains the report template used to organized the set of valid and relevant sentences to generate dynamic self-tracking reports. In Section 5, the experimental setup is described, where the experiments and the obtained results are shown. Finally, Section 6 draws some conclusions and presents related future works.

2. General architecture

The general architecture is based on our research in the field of Computational Theory of Perceptions (CTP). CTP was introduced in Zadeh's seminal paper "From computing with numbers to computing with words - from manipulation of measurements to manipulation of perceptions" [31] and further developed in subsequent papers [32, 33]. It grounds on the fact that human cognition is based on the role of perceptions, and the remarkable capability to granulate information in order to perform physical and mental tasks without any traditional measurements and computations.

In this paper, we extend our previous architecture by adding two new processing modules: a Relevance Module that automatically distinguishes between the relevant and the irrelevant information according to the user's needs, and a Dynamic Expression Module that dynamically generates linguistic reports. In Fig. 1, we can identify two main stages for the linguistic description of time series: an off-line design process, where two different data structures are designed, and an on-line instantiation process, where the different processing modules are feed with time series data.

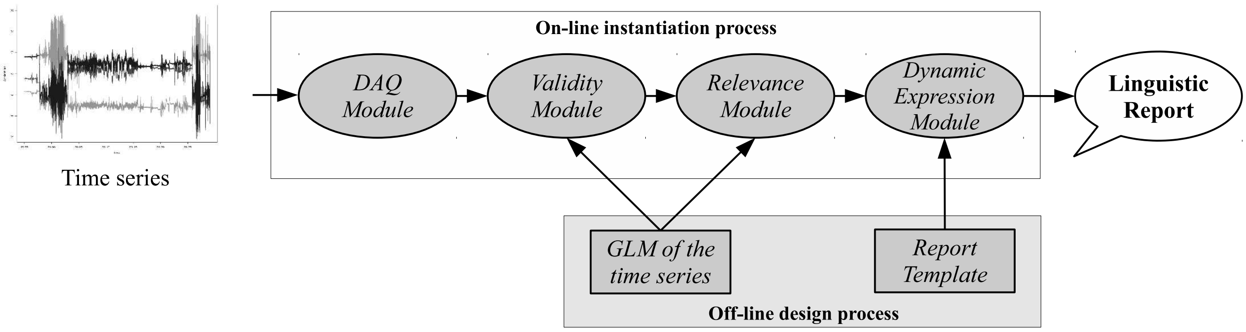


Figure 1: Architecture for generating dynamic reports for linguistic description of time series.

During the off-line design process, the designer collects and organizes a corpus of natural language (NL) expressions that are typically used in the application domain to describe the relevant features of the monitored time series. The designer analyzes the particular meaning of each linguistic expression in specific situation types to design a first data structure, which consists of the Granular Linguistic Model (GLM) of this time series. According to the user's requirements, the designer creates also another data structure consisting of the Report Template, which is used to organize the final report in a clear and orderly manner.

During the on-line instantiation process, the computational system feeds the Data Acquisition (DAQ) Module with the input time series in order to calculate the validity and relevance of each NL expression by means of the Validity Module and the Relevance Module, respectively. Finally, the Dynamic Expression Module creates a linguistic report by choosing and connecting those valid and relevant NL expressions and following the structure presented by the Report Template. We describe each module of this on-line instantiation process as follows:

- The Data Acquisition (DAQ) Module collects the time series either form sensors or by accessing to the information stored in a database. In this application context, the DAQ Module obtains numerical data from the triaxial accelerometer embedded in current smartphones. We have developed an experimental application based in the Android platform that acquires and stores the accelerations produced during the daily activities of a person. Current smartphones, despite not being specialized sensors, are commercial devices that are carried by millions of users. We exploit their great range of communications capabilities, integrated hardware and software features.

  The DAQ Module uses as input vector the time series of the accelerations (a_x, a_y, a_z) in m/s^2. It processes these input data to obtain the accelerations module ρ (Eq. 1) and its standard deviation σ using an experimentally obtained 20 seconds moving window (Eq. 2):

$$\rho = \sqrt{a_x^2 + a_y^2 + a_z^2} \tag{1}$$

$$\sigma = \sqrt{\frac{\sum_{i=1}^{N}(\rho_i - \bar{\rho})^2}{N-1}} \tag{2}$$

where, N is the number of data points in each 20 second moving window, ρ_i is each sample of the accelerations module and $\bar{\rho}$ is the average value. It is worth remarking that ρ and σ are invariant with respect to the smartphone orientation and position. We do not mind how the movement accelerations are distributed among x, y and z axes, but we are interested in knowing the magnitude and variability of the combined signals, which is represented by these two variables. This deals with one of the most important limitations of existing activity recognition systems, which require that the acquisition devices be placed in specific locations, such as trousers pocket, breast pocket or waist. Our system, therefore, estimates the level of physical activity by wearing the smartphone wherever the user wants.

- The Validity Module takes the numerical inputs acquired by the DAQ Module and instantiates the validity degrees of the NL expressions contained in the GLM in order to describe the phenomenon (the details of the GLM are explained in Section 3).

- After being instantiated with a set of input data, the GLM provides a structure of valid sentences that in medium size applications could include hundreds of sentences. Therefore, it is critical to perform a relevance analysis in order to select and collect the relevant sentences within a document that highlights the interesting characteristics of the time series.

  In this paper, we propose a way of automatically calculating the relevance values of each sentence. When we want to automatically generate a linguistic report similar to the obtained by an expert, it is difficult to determine, a priori, which information should appear in it. Usually, the content of reports depends on the obtained results, and the report generator needs to be able to assign the relative importance of each perception to each situation type.

- The Dynamic Expression Module, based on the structure presented by the Report Template, creates the linguistic report during the on-line instantiation process by choosing and connecting the adequate NL expressions. Once the relevance and the validity of each possible NL expression are calculated by the previous modules, the Dynamic Expression Module combines those sentences that accumulate the highest values by calculating the product of the validity and the relevance values of each NL expression. Therefore, depending on input data, different reports will be obtained each time, since they depend on the validity and specially on the automatically calculated relevance values.

3. Granular Linguistic Model to self-track the physical activity

The main element of this structure is known as Computational Perception (CP), which is based on the concept of linguistic variable developed by Zadeh [34]. CPs are computational models of units of information (granules) acquired by the designer about the time series to be modeled at certain granularity degree. A CP is a tuple with three components (A, W, R) described as follows:

A is a multidimensional matrix of linguistic expressions (words or sentences in NL) that represents the whole linguistic domain of CP. During the design stage, these values are defined by the designer extracting suitable sentences from the linguistic corpus of the application domain.

W is a multidimensional matrix of validity degrees whose elements are in the interval $[0, 1]$ assigned to each element of A in a specific context. The validity value represents the precision of each sentence to describe the specific input data. During the on-line instantiation process, these values will be dynamically assigned by the Validity Module reflecting the current state of the monitored time series.

R is a multidimensional matrix of relevance degrees whose elements are also in the interval $[0, 1]$ assigned to each element of A in a specific context. The values of relevance depends on the application. In

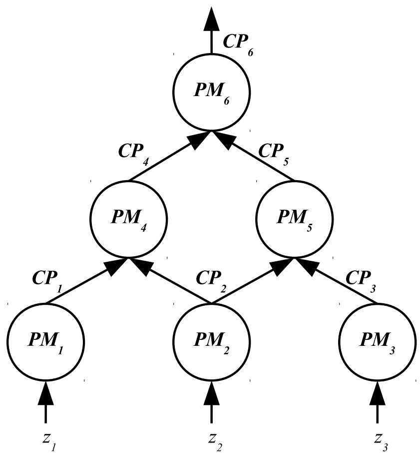


Figure 2: Example of a GLMP.

simple applications, they can be assigned by the designer during the off-line design stage. However, complex applications such as the presented in this paper, demand relevance matrices that will be calculated by the Relevance Module during the on-line instantiation process based on the validity degrees of higher or lower CPs, as we will show later.

A GLM consists of a network of Perception Mappings (PMs) used to create and aggregate CPs. Fig. 2 shows an example of GLM of a phenomenon. In this example, the phenomenon can be described at a very basic level in terms of three numeric variables provided by values z_1, z_2, and z_3, respectively. Then, other two higher-level descriptions of the phenomenon are provided. These descriptions are given in the form of Computational Perceptions CP_4 and CP_5. The Perception Mappings PM_4 and PM_5 explain CP_4 and CP_5 in terms of CP_1, CP_2, and CP_3. Finally, the top-order description of the phenomenon is provided, at the highest level of abstraction, by CP_6, explained by PM_6 in terms of CP_4 and CP_5. Notice that, using this structure, one can provide not only a linguistic description of the phenomenon at certain level, but an explanation in terms of linguistic expressions at a lower level. We will see how, after being instantiated with a set of input data, the GLM provides a structure of valid sentences that in medium size applications could include hundreds of sentences.

In the application case of this paper, the GLM provides three different granularity levels that describe the phenomenon. Each granularity level is composed by CPs that describe the phenomenon by means of instant, daily or weekly information. In dashed lines, the GLM represents the relationship among PMs to calculate the relevance degrees, and in continuous lines the relationship to calculate the validity degrees. The following subsections detail each PM included in the GLM developed to self-track the physical activity of a person by analyzing his/her basic movements, such as walking, running and stopped (Fig. 3).

3.1. Standard deviation of accelerations module (PM_σ)

A PM is a tuple (U, y, f, g, T) where each component is explained as follows:

U is a set of numerical values or input CPs. Here, it is composed by the time series of the standard deviation (σ) of the accelerations module.

y is the output $CP_\sigma = (A_\sigma, W_\sigma, R_\sigma)$, where $A_\sigma =$ (*Low (L), Medium (M), High (H)*).

f is the validity function $W_\sigma = f(\sigma)$ obtained by means of trapezoidal membership functions forming a strong fuzzy partition. At first, the linguistic labels of A_σ were uniformly distributed but we have analyzed the statistical distribution of this variable throughout the acquired experimental data. Thus, these labels were heuristically defined by their vertices as follows: $\{$*Low* $(0, 0, 0.001, 0.015)$, *Medium* $(0.001, 0.015, 0.08, 0.095)$, *High* $(0.08, 0.095, \infty, \infty)\}$.

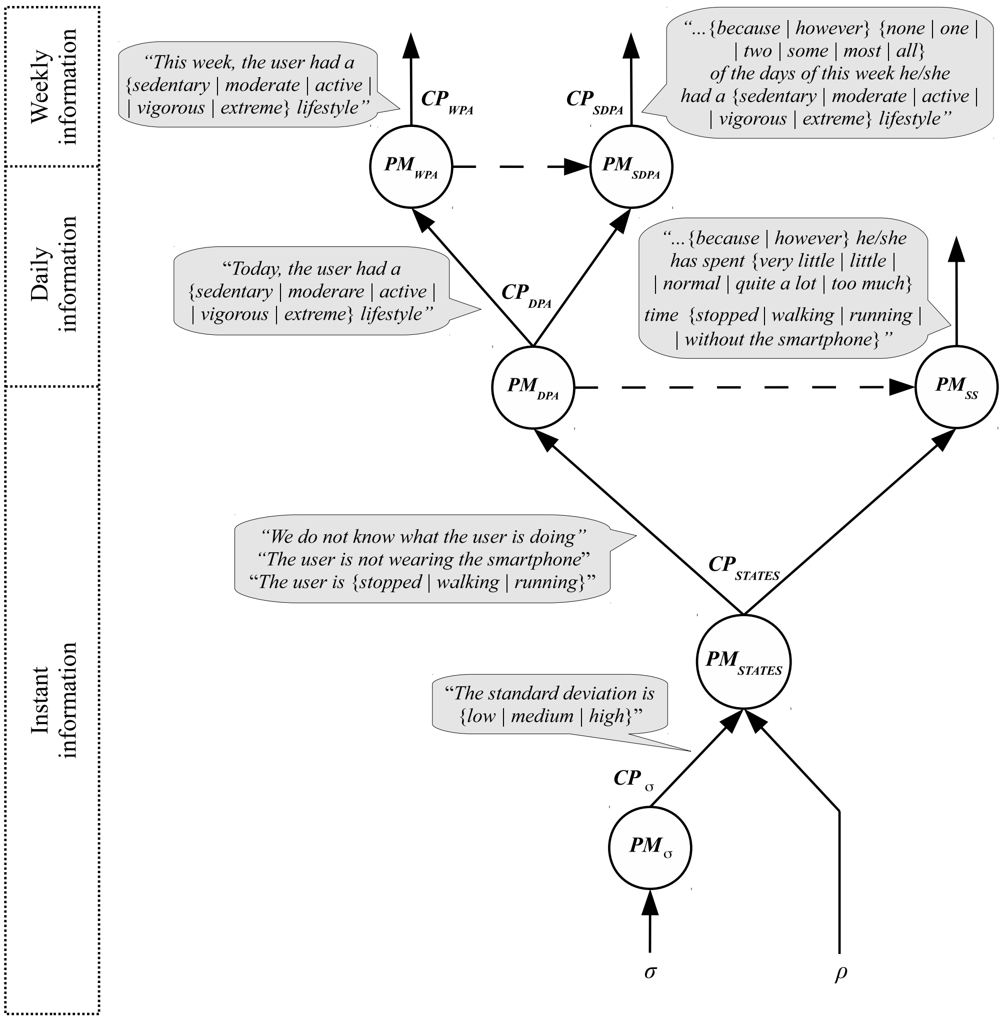


Figure 3: GLM of the time series that linguistically describes the physical activity.

g is the function that calculates the relevance of each linguistic label of A_σ to get the relevance matrix. In this case, since this CP works only as input data and it is not used in the linguistic reports, the relevance values are set to zero by the designer during the off-line design stage, $R_\sigma = (0, 0, 0)$. Therefore, the information of this CP will never appear in the final report.

T is a basic text generation algorithm, a template, that produces linguistic expressions as follows: *"The standard deviation is $\{low \mid medium \mid high\}$"*.

As a way to illustrate the computational process generated in a PM, we are going to analyze in detail the components explained above for a particular example of input data. In this PM_σ, U is a set of numerical values corresponding to the standard deviation of accelerations module. Provided that, at timestamp t, $U[t] = 0.01$. Then, according to the linguistic labels of A_σ, the validity function f determines that the input value is 0.95 *low* and 0.05 *medium*, i.e., $W_\sigma = (0.95, 0.05, 0)$. Note that here, g does not calculate the relevance of this information in the final report, since it is predefined by the designer during the off-line stage ($R_\sigma = (0, 0, 0)$). Finally, the text generation algorithm T produces the expression *"The standard deviation is low"*, with a validity degree of 0.95. In order to be aggregated and combined with other CPs or input numerical data, all this information is transmitted upstream.

3.2. States of physical activity (PM_{STATES})

In this perception mapping, we identify three of the basic physical activities established in [35]: *stopped*, *walking* and *running*. In addition, we define other two possible states, namely, the *unknown* and the *idle* states. The *unknown* state deals with representing the non-interpretability of the signal, providing high robustness to the system (for more information about uninterpretable data see [36]). It also represents those situations in which data is not captured, e.g., when the smartphone's sensor is malfunctioning or the device is switched off. The *idle* state represents those instants in which the user is not wearing the smartphone, e.g., when he/she leaves it on a desk. This PM has the following elements:

U is a vector composed by two inputs: the CP_σ and the numerical input data corresponding to the accelerations module (ρ).

y is the output CP_{STATES}, where A_{STATES} = (*Unknown* (q_0), *Idle* (q_1), *Stopped* (q_2), *Walking* (q_3), *Running* (q_4)).

f is the aggregation function that calculates the matrix of validity degrees $W_y = f(W_1, W_2, \ldots, W_n)$, where W_i are the matrices of validity degrees of the n input CPs. Here, $W_{STATES} = f(W_\sigma, \rho)$, where f calculates the validity degrees of the output CP_{STATES}. We use a Fuzzy Finite State Machine (FFSM) to implement this aggregation function. For a more detailed description of the FFSM paradigm and its applications, the interested reader could see [37].

According to the states diagram shown in Fig. 4, f is composed by rules R_{ii} to remain in the state q_i and rules R_{ij} to change from the state q_i to the state q_j. These fuzzy rules have the following structure:

$$R_{ii}: \text{IF } (State[t] \text{ is } q_i) \text{ AND } (\textit{Input variables constraints})_{ii} \text{ AND } (\textit{Temporal constraints})_{ii} \text{ THEN } (State[t+1] \text{ is } q_i) \tag{3}$$

$$R_{ij}: \text{IF } (State[t] \text{ is } q_i) \text{ AND } (\textit{Input variables constraints})_{ij} \text{ AND } (\textit{Temporal constraints})_{ij} \text{ THEN } (State[t+1] \text{ is } q_j) \tag{4}$$

The antecedent of each rule is composed by the state of physical activity at timestamp t, the input variables constraints, and the temporal constraints corresponding to that state, which are described below in Section 3.2.1 and Section 3.2.2, respectively. In Section 3.2.3, we detail the set of fuzzy rules.

g is the aggregation function that calculates the matrix of relevance degrees $R_y = g(W_1, W_2, \ldots, W_m)$, where W_i are the matrices of validity degrees of the m input CPs. In this case, this information will be always shown in the instant reports, therefore the relevance values are set to one by the designer during the off-line design stage, $R_{STATES} = (1, 1, 1, 1, 1)$.

T produces linguistic expressions that can be adapted depending on the situation. When the model is not able to recognize the input signal (q_0), the system reports the following expression: *"We do not know what the user is doing"*. However, when the system detects that the user does not carry the smartphone (q_1), the report is: *"The user is not carrying the smartphone"*. Finally, when the system is in the states q_2, q_3 and q_4, the template is: *"The user is* {*stopped* | *walking* | *running*}*."*.

3.2.1. Constraints on the input variables

The input variable constraints are defined over the standard deviation of the accelerations module (CP_σ) and the cadence of the periodical movement when the user is walking or running. This cadence, i.e., the period, is obtained by applying the Fast Fourier Transform (FFT) to the accelerations module (ρ) during a temporal moving window when CP_σ is *high*, i.e., the user is moving. The cadence is defined as follows: {*low* $(0, 0, 130, 135)$, *high* $(130, 135, \infty, \infty)$}. See in 3.2.3 how this constraints are used to define the antecedents of the fuzzy rules.

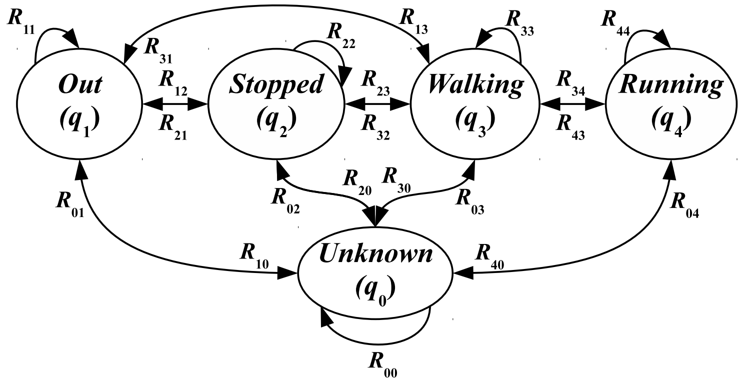


Figure 4: States diagram of the FFSM.

3.2.2. Temporal constraints

To provide robustness to the system, we have introduced temporal constraints over the length of the states. We numerically compute the duration d_{q_j} as the time that the constraints on the input variables of q_j are fulfilled, increasing d_{q_j} according to the sampling rate.

Temporal fuzzy constraints represent the minimum duration (in seconds) that the input variables have to be stable to enable the change from state q_i to state q_j. Thus, we have defined the linguistic labels *enough* and *not enough* to indicate when the duration d_{q_j} is enough or not to consider the change to the state q_j. Each possible change of state has its own definition for these two linguistic labels: d_{Idle} is $\{$*enough* $(-0, 0, 290, 310)$, *not enough* $(290, 310, \infty, \infty)\}$, $d_{Stopped}$ is $\{$*enough* $(0, 0, 20, 40)$, *not enough* $(20, 40, \infty, \infty)\}$, $d_{Walking}$ is $\{$*enough* $(0, 0, 10, 15)$, *not enough* $(10, 15, \infty, \infty)\}$, and $d_{Running}$ is $\{$*enough* $(0, 0, 10, 15)$, *not enough* $(10, 15, \infty, \infty)\}$. These labels have been heuristically defined since the definition of very small durations could produce continuous changes among states (unstable model) and the definition of very large durations could produce a very static model.

3.2.3. Set of fuzzy rules

According to Eqs. 3 and 4, the rule base that models the physical activity is the following:

R_{01}: IF (*State*$[t]$ is *Unknown*) AND (σ is *low*) AND (d_{Idle} is *enough*) THEN (*State*$[t+1]$ is *Idle*)

R_{11}: IF (*State*$[t]$ is *Idle*) AND $\Big($(σ is *low*) OR (σ is *medium* AND $d_{Stopped}$ is *not enough*) OR (σ is *high* AND $d_{Walking}$ is *not enough*)$\Big)$ THEN (*State*$[t+1]$ is *Idle*)

R_{21}: IF (*State*$[t]$ is *Stopped*) AND (σ is *low*) AND (d_{Idle} is *enough*) THEN (*State*$[t+1]$ is *Idle*)

R_{31}: IF (*State*$[t]$ is *Walking*) AND (σ is *low*) AND (d_{Idle} is *enough*) THEN (*State*$[t+1]$ is *Idle*)

R_{02}: IF (*State*$[t]$ is *Unknown*) AND (σ is *medium*) AND ($d_{Stopped}$ is *enough*) THEN (*State*$[t+1]$ is *Stopped*)

R_{12}: IF ($S[t]$ is *Idle*) AND (σ is *medium*) AND ($d_{Stopped}$ is *enough*) THEN (*State*$[t+1]$ is *Stopped*)

R_{22}: IF (*State*$[t]$ is *Stopped*) AND $\Big($(σ is *low* AND d_{Idle} is *not enough*) OR (σ is *medium*) OR (σ is *high* AND $d_{Walking}$ is *not enough*)$\Big)$ THEN (*State*$[t+1]$ is *Stopped*)

R_{32}: IF (*State*$[t]$ is *Walking*) AND (σ is *medium*) AND ($d_{Stopped}$ is *enough*) THEN (*State*$[t+1]$ is *Stopped*)

R_{03}: IF (*State*$[t]$ is *Unknown*) AND (σ is *high* AND FFT_ρ is *low* AND $d_{Walking}$ is *enough*) THEN (*State*$[t+1]$ is *Walking*)

R_{13}: IF (*State*[*t*] is *Idle*) AND (σ is *high* AND FFT_ρ is *low* AND $d_{Walking}$ is *enough*) THEN (*State*[$t+1$] is *Walking*)

R_{23}: IF (*State*[*t*] is *Stopped*) AND (σ is *high* AND FFT_ρ is *low* AND $d_{Walking}$ is *enough*) THEN (*State*[$t+1$] is *Walking*)

R_{33}: IF (*State*[*t*] is *Walking*) AND ((σ is *low* AND d_{Idle} is *not enough*) OR (σ is *medium* AND $d_{Stopped}$ is *not enough*) OR (σ is *high* AND FFT_ρ is *low*)) THEN (*State*[$t+1$] is *Walking*)

R_{43}: IF (*State*[*t*] is *Running*) AND (σ is *high* AND FFT_ρ is *low* AND $d_{Walking}$ is *enough*) THEN (*State*[$t+1$] is *Walking*)

R_{04}: IF (*State*[*t*] is *Unknown*) AND (σ is *high* AND FFT_ρ is *high* AND $d_{Running}$ is *enough*) THEN (*State*[$t+1$] is *Running*)

R_{34}: IF (*State*[*t*] is *Walking*) AND (σ is *high* AND FFT_ρ is *high* AND $d_{Running}$ is *enough*) THEN (*State*[$t+1$] is *Running*)

R_{44}: IF (*State*[*t*] is *Running*) AND (σ is *high* AND FFT_ρ is *high* AND $d_{Running}$ is *enough*) THEN (*State*[$t+1$] is *Running*)

R_{i0}: ELSE (*State*[$t+1$] is *Unknown*)

The validity degrees of the sentences associated to each specific state of physical activity are calculated as a weighted average of the individual rules, where the weight of each rule R_{ij} corresponds to its firing degree τ_{ij}. This firing degree is calculated using the minimum for the AND operator and the bounded sum of Łukasiewicz [38] for the OR operator. We chose the state *Unknown* (q_0) as initial state of the time series, having a validity degree $w_{q_0} = 1$. The validity degree of the state q_0 is obtained by means of Eq. 5:

$$w_{q_0}[t+1] = \begin{cases} 1 - \sum\limits_{i=0}^{n}\sum\limits_{j=1}^{n}\tau_{ij} & \text{if} \quad \sum\limits_{i=0}^{n}\sum\limits_{j=1}^{n}\tau_{ij} \leq 1 \\ 0 & \text{if} \quad 1 < \sum\limits_{i=0}^{n}\sum\limits_{j=1}^{n}\tau_{ij} \end{cases} \tag{5}$$

The validity degree of the rest of states is calculated by means of Eq. 6:

$$w_{q_j}[t+1] = \begin{cases} \sum\limits_{i=0}^{n}\tau_{ij} & \text{if} \quad \sum\limits_{i=0}^{n}\sum\limits_{j=1}^{n}\tau_{ij} \leq 1 \\ \dfrac{\sum_{i=0}^{n}\tau_{ij}}{\sum_{i=0}^{n}\sum_{j=1}^{n}\tau_{ij}} & \text{if} \quad 1 < \sum\limits_{i=0}^{n}\sum\limits_{j=1}^{n}\tau_{ij} \end{cases} \tag{6}$$

3.3. Daily physical activity (PM_{DPA})

The interest of standard references for estimating the level of physical activity is growing and the need for reliable, objective and portable energy measurement systems is evident. We define the level of physical activity as a measure that depends on the duration, frequency and intensity of the physical movements performed by a person during a period of time, typically one day. This PM estimates the user's level of daily physical activity (DPA). It has the following elements:

U is the input CP_{STATES}.

y is the output CP_{DPA}, where A_{DPA} = (*Sedentary (S), Moderate (M), Active (A), Vigorous (V), Extreme (E)*).

f is the validity function that calculates the different validity degrees of the NL expressions associated to the user's daily physical activity based on its numerical value, which is calculated using Eq. 7:

$$\mathrm{DPA} = \frac{\sum_{i=0}^{4} TD_i \cdot UW \cdot AF \cdot GF}{t} \tag{7}$$

where TD_i is the total duration that the user has practiced the activity i (0=*unknown*, 1=*idle*, 2=*stopped*, 3=*walking* and 4=*running*) during the recording period t (in minutes), UW is the user's weight expressed in kilograms and GF is a gender factor that takes value 1 if the user is a man and 0.9 if she is a woman. AF is a factor whose value depends on the activity as follows: 0.0285 (*unknown*, *idle* and *stopped*), 0.063 (*walking*) and 0.2 (*running*). This way to estimate the physical activity has been obtained from [35], where an analogous crisp way is proposed to estimate the energy consumption. During the *unknown* or *idle* states, we consider that the level of physical activity is, at least, equal to the minimal level of physical activity, that it is produced when he/she is *stopped*.

As seen in Eq. 7, the levels of DPA vary depending on the gender and weight of the user. Therefore, the distribution of the linguistic labels that calculate the user's lifestyle based on the variable DPA has to be tuned according to the user's gender and weight. Therefore, the vertices that define the linguistic labels depend on a scale factor C_1, which is calculated using Eq. 8: $\{S\ (0, 0, 2.23C_1, 2.37C_1),\ M\ (2.23C_1, 2.37C_1, 2.58C_1, 2.72C_1),\ A\ (2.58C_1, 2.72C_1, 2.93C_1, 3.07C_1),\ V\ (2.93C_1, 3.07C_1, 3.28C_1, 3.42C_1),\ E\ (3.28C_1, 3.42C_1, \infty, \infty)\}$, where

$$C_1 = \frac{UW \cdot GF}{80} \tag{8}$$

UW is the user's weight (in kilograms), GF is the gender factor explained above, and 80 is the weight of the user taken as reference.

g is the function that calculates the relevance of each linguistic label of A_{DPA} to get the relevance matrix. In this case, this information will be always shown in the daily reports, therefore the relevance values are set to one by the designer during the off-line design stage, $R_{DPA} = (1, 1, 1, 1, 1)$.

T produces linguistic expressions as follows: *"Today, the user had a {sedentary | moderate | active | vigorous | extreme} lifestyle"*.

3.4. Daily summary of states (PM_{SS})

This PM complements the information provided by CP_{DPA}, describing the amount of samples in which the user has not been wearing the smartphone (*idle* state), *stopped*, *walking* or *running* during the day. Thanks to the sampling rate, the aggregation of the number of samples directly gives us an idea about the time that the user has been in each state. Thus, we can convert, e.g. *"a small number of samples"* into *"little time"*. For example, if the daily physical activity was *vigorous*, this PM obtains sentences that complements this information, e.g., *"...because he/she has spent quite a lot time running"*. This PM has the following elements:

U is the input CP_{STATES}.

y is the output CP_{SS}, where A_{SS} = (*Very small (VS), Small (S), Normal (N), Quite a lot (Q), Too much (TM)*).

f is the aggregation function that uses the $\alpha-$cuts method proposed in [39] to calculate quantifiers. This method is appropriate to deal with some well-known problems present in other methods of calculating quantifiers. We have successfully applied it in previous research works [25]. The procedure is as follows:

The first step is calculating the percentage of samples labeled as *idle*, *stopped*, *walking* and *running*, that are contained at each $\alpha-$level ($N_{\alpha j}$), where j determines the index of each state. We calculate it over the total amount of samples (n) of the day (from 9 a.m. to 9 p.m.), by means of Eq. 9, being

$\alpha \in \mathrm{K} = \{0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1\}$. The number of elements in the set K is the selected level of resolution, i.e., $|\mathrm{K}| = 10$ in this particular case.

$$N_{\alpha j} = \Big(\frac{1}{n}\sum_{i=1}^{n} F_\alpha(w_{q_j}[i])\Big) \times 100 \tag{9}$$

where:

$$F_\alpha(w_{q_j}) = \begin{cases} 1 & \text{if} \quad w_{q_j} \geq \alpha \\ 0 & \text{if} \quad w_{q_j} < \alpha \end{cases} \tag{10}$$

Then, we calculate the membership degree of each $N_{\alpha j}$ with the set of linguistic labels of A_{SS}. In this paper, as an example case, we have defined the shape of each linguistic label as specified in Table 1. Note that these shapes are determined by the physical activity that they represent, e.g., the percentage of samples corresponding to *"too much time running"* are smaller than the ones corresponding to *stopped* or *walking*. When the application case consists of monitoring the physical evolution of patients in medical treatments or medical interventions, this definition has to be done by an expert according to the final goal .

	Very small	*Small*	*Normal*	*Quite a lot*	*Too much*
Idle	$(0, 0, 5, 10)$	$(5, 10, 30, 35)$	$(30, 35, 55, 60)$	$(55, 60, 80, 85)$	$(80, 85, 100, 100)$
Stopped	$(0, 0, 0, 5)$	$(0, 5, 35, 40)$	$(35, 40, 55, 60)$	$(55, 60, 75, 80)$	$(75, 80, 100, 100)$
Walking	$(0, 0, 5, 10)$	$(5, 10, 25, 30)$	$(25, 30, 45, 50)$	$(45, 50, 75, 80)$	$(75, 80, 100, 100)$
Running	$(0, 0, 0, 2.5)$	$(0, 2.5, 5, 7.5)$	$(5, 7.5, 10, 12.5)$	$(10, 12.5, 15, 17.5)$	$(15, 17.5, 100, 100)$

Table 1: Definition of linguistic labels to summarize each physical activity.

Finally, the last step consists of obtaining each $w_{SS_{ij}}$ by calculating the average value of the contribution of each $\alpha-$level using Eq. 11.

$$w_{SS_{ij}} = \frac{1}{|\mathrm{K}|}\sum_{\forall\alpha\in\mathrm{K}} \mu_{Q_i}(N_{\alpha j}) \tag{11}$$

g is the function that calculates the relevance of each NL expression of A_{SS} in order to get the relevance matrix R_{SS}. In this case, as can be seen in Fig. 3 (see the dashed lines that show this relationship), the relevance of this output CP depends on the validity degrees of the sentences that calculate the daily physical activity (CP_{DPA}), i.e., $R_{SS} = g(W_{DPA})$. This is one of the most interesting points of our proposal, because the definition of this interaction among validity degrees of CP_{DPA} and relevance of CP_{SS} will generate dynamic reports that will be adapted to the user's needs. In this case, the relevance of CP_{SS} will depend of the validity degrees of CP_{DPA}.

	Because					However				
	VS	*S*	*N*	*Q*	*TM*	*VS*	*S*	*N*	*Q*	*TM*
Idle	0	0	0	0	0	0	0	0	1	1
Stopped	$A \cup V \cup E$	$A \cup V \cup E$	M	$S \cup M$	S	0	0	0	0	0
Walking	$S \cup M$	A	$A \cup V$	$A \cup V \cup E$	$V \cup E$	0	$S \cup M$	0	0	0
Running	0	A	$A \cup V$	$A \cup V \cup E$	$V \cup E$	0	$S \cup M$	0	0	0

Table 2: Relevance matrices for daily summary of states.

In the presented application, the designer has decided that it is interesting to always include the information related to the daily lifestyle of users in the daily physical reports ($R_{DPA} = (1, 1, 1, 1, 1)$)

but not to always include all the information about the amount of time that the user has spent in a specific activity along the day. In Table 2, can be seen the structure of the rules that function g implements to calculate the relevance values. There are two different relevance groups, the first one contributes to the dynamic report by explaining the causes ("because" group) while the second one shows some untoward causes ("however" group). Each element of the table corresponds to an element of the relevance matrix and will be calculated in terms of the validity degrees of CP_{DPA}.

For example, in the "because" group, the relevance of the NL expression associated to the *running* state during quite a lot time ($r_{SS_{44}}$) will be obtained from the union (implemented as the bounded sum of Łukasiewicz denoted by Ł) of the validity degrees of an *active* (w_{DPA_3}), *vigorous* (w_{DPA_4}), or *extreme* (w_{DPA_5}) lifestyle during the day, i.e., $r_{SS_{44}} = \text{Ł}(w_{DPA_3}, w_{DPA_4}, w_{DPA_5})$.

Note that, there are also elements whose relevance is fixed by the designer and are independent of the validity degrees of CP_{DPA}. E.g., in the "however" group, an *idle* state that has been activated during quite a lot or too much time will have the highest relevance degree, denoted by 1 ($r_{SS_{14}} = r_{SS_{15}} = 1$). On the other hand, there are also NL expressions completely irrelevant and denoted by 0, such as the NL expression associated to the *running* state during quite a lot time ($r_{SS_{44}} = 0$).

T provides the following set of linguistic expressions: *"...{because | however} he/she has spent {very little | little | normal | quite a lot | too much} time {without the smartphone | stopped | walking | running}"*.

3.5. Weekly physical activity (PM_{WPA})

U is the vector of the numerical values of DPA.

y is the output CP_{WPA}, where A_{WPA} = (*Sedentary (S), Moderate (M), Active (A), Vigorous (V), Extreme (E)*).

f is the output function that calculates the validity degrees of the output CP_{WPA}. Here, we have aggregated the DPA corresponding to each day of the week by means of the weighted average represented by Eq. 12. Here, since we are interested in highlighting extreme behaviors like sedentary or hyperactive lifestyles, we have weighted them with higher coefficients. These coefficients can be tuned by the therapist in order to highlight other types of lifestyles according to each user pathology.

$$\text{WPA} = \frac{\sum_{i=1}^{7} K[i] \cdot DPA[i]}{\sum_{i=1}^{7} K[i]} \tag{12}$$

$$K[i] = 10 \cdot (w_{DPA_1}[i]) + 5 \cdot (w_{DPA_2}[i]) + (w_{DPA_3}[i]) + 5 \cdot (w_{DPA_4}[i]) + 10 \cdot (w_{DPA_5}[i])$$

where, $DPA[i]$ is the DPA corresponding to each day i and $(w_{DPA_1}[i]), (w_{DPA_2}[i]), \ldots$ are the validity degrees of the output CP_{DPA} calculated in Section 3.3. The shape of the linguistic labels is the same as the ones presented in that section.

g is the function that calculates the relevance of each linguistic label of A_{WPA} to get the relevance matrix. In this case, this information will be always shown in the weekly reports, therefore the relevance values are set to one by the designer during the off-line design stage, $R_{WPA} = (1, 1, 1, 1, 1)$.

T produces linguistic expressions as follows: *"This week, the user had a {sedentary | moderate | active | vigorous | extreme} lifestyle"*.

3.6. Weekly summary of DPAs (PM_{SDPA})

This PM complements the information provided by CP_{WPA}, describing the amount of days (*None (N), One (O), Two (T), Some (S), Most (M), All (A)*) that the user had a *Sedentary (S)*, *Moderate (M)*, *Active (A)*, *Vigorous (V)* or *Extreme (E)* lifestyle during the week. For example, giving that the weekly physical activity was *vigorous*, this PM obtains sentences that complements this information, e.g., *"...however, two of the days of this week he/she had a moderate lifestyle"*. It has the following elements:

U is the input CP_{DPA}.

y is the output CP_{SDPA}, where A_{SDPA} is a bidimensional matrix of size 6×5:

$$A_{SDPA} = \begin{pmatrix} None(N)/Sedentary(S) & One(O)/Sedentary(S) & \cdots & All(A)/Sedentary(S) \\ None(N)/Moderate(M) & One(O)/Moderate(M) & \cdots & All(A)/Moderate(M) \\ \vdots & \vdots & \ddots & \vdots \\ None(N)/Extreme(E) & One(O)/Extreme(E) & \cdots & All(A)/Extreme(E) \end{pmatrix}$$

f is the function that aggregates the daily physical activity over the whole week. Here, rather than using the α-cuts method, for the sake of simplicity and in order to provide different examples of aggregation functions, we use the fuzzy quantifiers of Zadeh [40]. Here, the shapes of the linguistic quantifiers are defined by their vertices as follows:{*None* $(-\infty, -\infty, 0, 1)$, *One* $(0, 1, 1, 2)$, *Two* $(1, 2, 2, 3)$, *Some* $(2, 3, 4, 5)$, *Most* $(4, 5, 6, 7)$, *All* $(7, 7, \infty, \infty)$}. The validity degree of each quantifier n for each lifestyle m is calculated using Eq. 13.

$$w_{SDPA_{mn}} = \frac{\sum_{d=1}^{7} w_{DPA_m}[d]}{7} \tag{13}$$

where d is the day of the week.

g is the function that, in order to get the relevance matrix R_{SDPA}, calculates the relevance of each NL expression of A_{SDPA}. In this case, as can be seen in Fig. 3, the relevance of this CP depends on the validity degrees of the sentences that calculate the weekly physical activity (CP_{WPA}).

The relevance matrix deals with explaining CP_{WPA} in terms of CP_{SDPA}, i.e., the relevance of the sentences of CP_{SDPA} will depend of the validity degrees of CP_{WPA}: $R_{SDPA} = g(W_{WPA})$.

	Because						However					
	None	*One*	*Two*	*Some*	*Most*	*All*	*None*	*One*	*Two*	*Some*	*Most*	*All*
Sedentary	0	$S \cup M$	$S \cup M$	0	S	S	0	A	A	0	0	0
Moderate	0	0	A	$A \cup M$	M	M	0	V	V	0	0	0
Active	0	0	0	A	A	A	0	0	0	0	0	0
Vigorous	0	0	A	$A \cup V$	V	V	0	M	M	0	0	0
Extreme	0	$E \cup V$	$E \cup V$	0	E	E	0	A	A	0	0	0

Table 3: Relevance matrix for weekly summary of physical activity.

In Table 3, can be seen the structure of the rules that function g implements to calculate the relevance matrix. Each element of the table corresponds to an element of the relevance matrix and will be calculated in terms of the validity degrees of CP_{WPA}. Here, there are also two different groups of relevance, the first group contributes to the dynamical report template by explaining the causes ("because" group) while the second one shows some untoward causes ("however" group).

T produces the following linguistic expressions: *"...{because | however} {none | one | two | some | most | all} of the days of this week, he/she had a {sedentary | moderate | active | vigorous | extreme} lifestyle".*

3.7. Remarks

Throughout this section, we have presented a set of parameters whose definition will modify the final results. In order to collect them and improve the understanding of the paper, we have elaborated the Table 4. This table presents the set of parameters and classifies them into three different types, according to the following criteria:

- Parameters of type 1 are defined by the designer during the off-line stage and belong to the generic model of physical activity, regardless of the end user or the application.
- Parameters of type 2 are also defined by the designer during the off-line stage but, in this case, they depend on the application use. For example, these parameters could change if we develop applications either to analyze the activity of very athletic people or we are focused on elderly with reduced mobility.
- Finally, parameters of type 3 depend on the specific end user characteristics. Typically, the definition of these parameters needs the collaboration of experts in the field, tuning them by means of the adequate human-computer interface.

Table 4 classifies and describes the adjustable parameters of this paper:

Parameter	Type	Comment
A_σ linguistic labels definition	2	These linguistic labels were heuristically defined. The correct definition of σ allows us to identify among different physical states (stopped, walking, running, ...). Depending on the situation, this definition may need to be tunned, for example, when monitoring the physical activity of someone with highly reduced mobility such as elderly, decreasing the threshold to consider he/she is walking.
Temporal constraints in PM_{STATES}	1	Temporal constraints definition is used to control the minimum duration that the input conditions have to be stable to consider a change among physical states. They control the "flexibility" of the model to allow more or less changes among states. It makes sense to generate a stable model that omits small and punctual activities.
A_{SS} linguistic labels definition	3	These linguistic labels summarize the amount of time that the user has been in each identified activity, represented in terms of *very little, little, normal, quite a lot* or *too much* time. When the application monitors the physical evolution of users involved in medical treatments, this definition has to be done by experts, according to the final goal of the therapy.

Table 4: Identification of adjustable parameters.

4. Report Template to self-track the physical activity

The Report Template is created by the designer during the off-line design process. The GLM of the physical activity provides, associated to each CP, a collection of valid and relevant sentences that describe the user's lifestyle with different levels of temporal detail. The set of linguistic expressions obtained by the text generation algorithm (T) of each CP is just one of the possible options to describe the phenomenon. Indeed, this set of expressions is aimed to make the GLM as interpretable as possible for the designer and his/her work team.

In this application, the Report Template is able to provide the linguistic reports at different levels of temporal granularity (Fig. 5). First of all, it is able to provide instant reports, i.e., reports about the activity that the user is developing in each specific moment (CP_{STATES}). This is accompanied by a graphical representation of the activities evolution along the analyzed day (see an example in Section 5.2.1).

Secondly, the presented application deals with allowing the user to self-track his/her daily and weekly physical activity. Daily reports inform about the user's lifestyle (CP_{DPA}) and summarize the amount of time that he/she has spent in each activity along the whole day (CP_{SS}). Weekly reports inform about the average user's lifestyle during the week (CP_{WPA}) and summarize the quantity of days that he/she has had each lifestyle (CP_{SDPA}). These daily and weekly reports will adopt the structure presented in Figs. 6

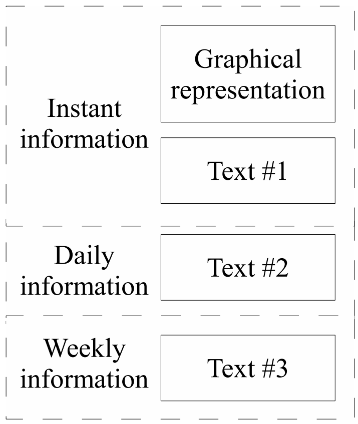


Figure 5: Report template to self-track the physical activity.

and 7, respectively. In this sense, as we have explained in Sections 3.3 and 3.5, the relevances of CP_{DPA} and CP_{WPA} are set to one ($R_{DPA} = R_{WPA} = (1,1,1,1,1)$), therefore this information will be always included in daily and weekly reports, respectively. On the other hand, the inclusion of CP_{SS} and CP_{SDPA} will be conditioned by the validity and relevance of the obtained results, calculated by f and g functions respectively. Thus, this information will appear, be partially omitted or fully omitted in the reports.

The Report Template includes beginning (START) and ending (END) messages. The beginning will vary depending on whether the report is instantaneous, daily or weekly. Thus, we can find expressions such as "*Currently,...*", "*On Monday,...*", and "*In summary, this week...*". On the other hand, the ending will include some additional information such as "*...which could condition the results*". In weekly reports, we use some linguistic techniques that provide the message with a more human-like content. We refer the interested reader to [41], for more details about this NL generation techniques. Some of these techniques consist on replacing the enumeration of all the weekday names by "*the rest of the week*" when we refer to all the days least one or two; saying "*weekend*" when we refer to Saturday and Sunday; and using expressions like "*decreased*" or "*increased*" when we compare the lifestyle of two periods. Finally, when the final report have to include two or more because/however sentences, they are linked by the appropriate nexus, e.g., punctuation marks and expressions like *"and"*, *"too"*, *"also"*, etc. If the final report have to include two "*however*" expressions, the second one is substituted by "*nevertheless*".

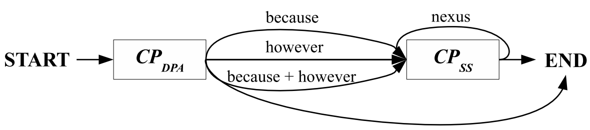


Figure 6: Report template for daily reports.

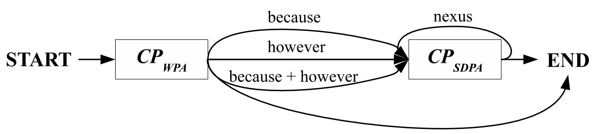


Figure 7: Report template for weekly reports.

The Dynamic Expression Module, according to the Report Template, allows the designer to analyze the GLM to highlight the most interesting and relevant features for users. Thanks to the potential of meaning of the developed GLM, users can manage a large amount of information in an adequate language domain, selecting the level of detail for each specific situation.

In order to choose the most adequate NL expression (or set of NL expressions) corresponding to each CP, the Dynamic Expression Module combines the validity and the relevance values of each expression by calculating the product of these two values.

As an example, consider how the Dynamic Expression Module chooses the NL expressions of the daily summary of states (CP_{SS}). In order to calculate which NL expression or set of NL expressions associated to the "because" group will be selected, the Dynamic Expression Module calculates the product of the validity (W_{SS}) and the relevance values (R_{SS}) as shown in Eq. 14:

$$
\begin{aligned}
because &= \begin{pmatrix} w_{SS_{11}} \cdot r_{SS_{11}} & \cdots & w_{SS_{15}} \cdot r_{SS_{15}} \\ w_{SS_{21}} \cdot r_{SS_{21}} & \cdots & w_{SS_{25}} \cdot r_{SS_{25}} \\ \vdots & \ddots & \vdots \\ w_{SS_{41}} \cdot r_{SS_{41}} & \cdots & w_{SS_{45}} \cdot r_{SS_{45}} \end{pmatrix} = \\
&= \begin{pmatrix} 0 & \cdots & 0 \\ w_{SS_{21}} \cdot \text{Ł}(w_{DPA_3}, w_{DPA_4}, w_{DPA_5}) & \cdots & w_{SS_{25}} \cdot w_{DPA_1} \\ \vdots & \ddots & \vdots \\ 0 & \cdots & w_{SS_{45}} \cdot \text{Ł}(w_{DPA_4}, w_{DPA_5}) \end{pmatrix}
\end{aligned} \tag{14}
$$

Similarly, to calculate which NL expression or set on NL expressions associated to the “however” group will be selected, the Dynamic Expression Module calculates the product of the validity and the relevance values as shown in Eq. 15:

$$
\begin{aligned}
however &= \begin{pmatrix} w_{SS_{11}} \cdot r_{SS_{11}} & \cdots & w_{SS_{15}} \cdot r_{SS_{15}} \\ w_{SS_{21}} \cdot r_{SS_{21}} & \cdots & w_{SS_{25}} \cdot r_{SS_{25}} \\ \vdots & \ddots & \vdots \\ w_{SS_{41}} \cdot r_{SS_{41}} & \cdots & w_{SS_{45}} \cdot r_{SS_{45}} \end{pmatrix} = \\
&= \begin{pmatrix} 0 & 0 & \cdots & 1 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & w_{SS_{42}} \cdot \text{Ł}(w_{DPA_1}, w_{DPA_2}) & \cdots & 0 \end{pmatrix}
\end{aligned} \tag{15}
$$

As stated in the Introduction, a good quality report must meet the four Grice’s Maxims. In this paper, each of these four maxims is considered in the linguistic descriptions as follows:

- We have designed an interpretable and transparent linguistic model that describes time series by providing a set of sentences with an associated validity degree (Maxim of Quality).
- It uses of a corpus of NL expressions typically used in the application domain. Thus, the model describes in everyday language the user’s lifestyle, so that everybody, experts and non-experts, can interpret the results (Maxim of Manner).
- The generated dynamic reports highlight the relevant information and omit the irrelevant (Maxim of Relation).
- Finally, results are described with different levels of detail, limiting the reports size. This tool generates reports with different abstraction levels according to the user’s needs, being from the most concise to the most general in each specific case (Maxim of Quantity).

5. Experimentation

5.1. Experimental layout

In order to verify the applicability and effectiveness of the developed computational tool, we performed an experiment based on tracking the physical activity of an 80 kilograms man for a period of five months (20 weeks). This subject performed daily activity self-reports that helped us to adjust the model. We obtained the three different types of reports to self-track the physical activity, i.e., instant reports that indicate in every moment what the user is doing; daily reports that linguistically describe his physical activity during the day; and weekly reports that summarize the user’s activity over the whole week.

The user started the application in the morning and stopped it at night, sending data to a server in order to be analyzed. We applied the GLM of the time series to the received data and estimated the user's instant, daily, and weekly physical activity as explained in Section 3.

5.2. Results

The following subsections present some instant, daily and weekly examples reports obtained from the experimentation:

5.2.1. Instant reports

Figure 8, shows a graphical representation of the instantaneous physical activity for a certain period of time. The CP_{STATES} provides at each time instant the activity that the user is developing among the five different states explained in Section 3.2. On the top of the figure, we can see the time series of the accelerations module (ρ) that was obtained from the user's smartphone. In the bottom, we can see the representation of the validity degree of each state (W_{STATES}).

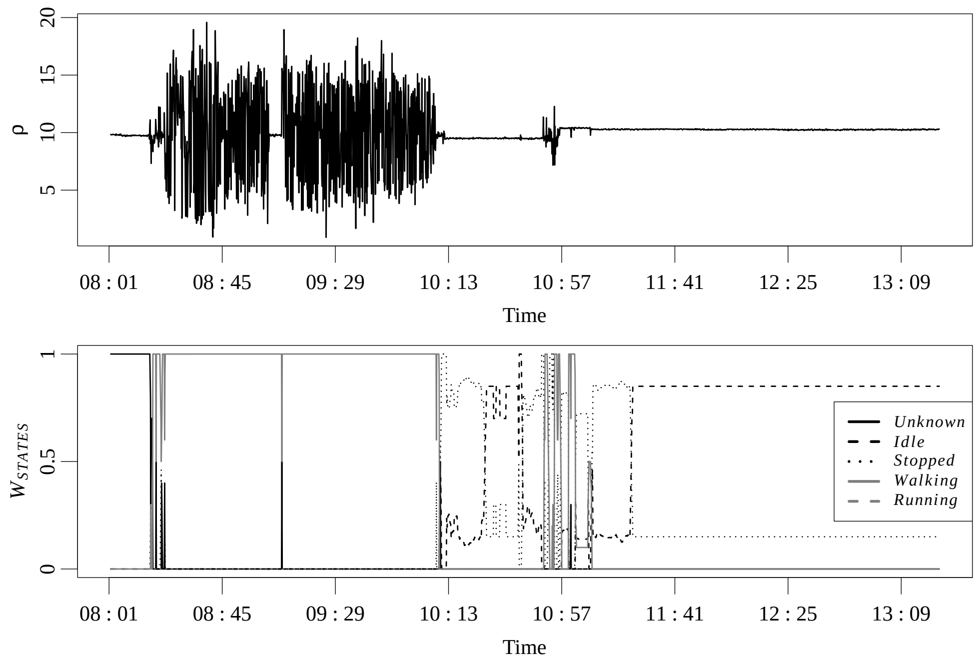


Figure 8: Instant report about the physical activity during the period of a day.

This graphic represents the physical activity performed from the beginning of the session to the instant report at 13:35 during the monitored day. It will be accompanied by the information of CP_{STATES}, which indicates the current activity state of the user:

"Currently, the user is not wearing the smartphone."

5.2.2. Daily reports

In order to demonstrate the potential and versatility of this concept, we have analyzed in detail one of the twenty monitored weeks. Table 5 lists the matrix of validity degrees (W_{SS}) associated to the daily physical activity corresponding to three consecutive days (Monday, Tuesday and Wednesday). We have added two different superscripts that represent those NL expressions that are relevant according to the relevance matrices calculated in the on-line instantiation stage by means of function g, that represent the "because" (B) and "however" (H) groups.

	Monday					Tuesday					Wednesday				
	VS	*S*	*N*	*Q*	*TM*	*VS*	*S*	*N*	*Q*	*TM*	*VS*	*S*	*N*	*Q*	*TM*
Idle	0	1	0	$\mathbf{0^H}$	$\mathbf{0^H}$	0	0	0.3	$\mathbf{0.7^H}$	$\mathbf{0^H}$	0	0	0	$\mathbf{1^H}$	$\mathbf{0^H}$
Stopped	$\mathbf{0^B}$	$\mathbf{0^B}$	0.8	0.2	0	$\mathbf{0^B}$	$\mathbf{0.8^B}$	0.2	0	0	0	1	$\mathbf{0^B}$	$\mathbf{0^B}$	0
Walking	0	$\mathbf{0.7^B}$	$\mathbf{0.3^B}$	$\mathbf{0^B}$	0	0	$\mathbf{0^B}$	$\mathbf{0.9^B}$	$\mathbf{0.1^B}$	0	$\mathbf{0^B}$	$\mathbf{1^H}$	0	0	0
Running	0.4	$\mathbf{0.6^B}$	$\mathbf{0^B}$	$\mathbf{0^B}$	0	0	$\mathbf{0^B}$	$\mathbf{0^B}$	$\mathbf{0^B}$	0	1	$\mathbf{0^H}$	0	0	0

Table 5: W_{SS} matrix corresponding to Monday, Tuesday and Wednesday physical activity.

On Monday of the selected week, the user had an *active* lifestyle ($w_{DPA_3} = 1$). Therefore, the relevance matrices of the "because" (B) and "however" (H) groups will be adapted according to the relevance matrix presented in Section 3.4 (Table 2). Following this table, the dynamic text report contains the following information:

"On Monday, the user had an active lifestyle, because he/she has spent little time walking and little time running."

On Tuesday, the user had also an *active* lifestyle ($w_{DPA_3} = 1$), which means that the calculated relevance matrix is the same that the one used on Monday. Here, we can observe how, despite having in both days an *active* lifestyle, the obtained daily reports are completely different, since the combination of W_{SS} and R_{SS}, where W_{SS} is different between Monday and Tuesday, gives different results. The dynamic text report contains the following information:

"On Tuesday, the user had an active lifestyle too, because he/she has spent a normal amount of time walking and little time stopped. However, he/she has spent quite a lot time in idle state, which could condition the results."

On Wednesday, the user's lifestyle decreases to *moderate* ($w_{DPA_2} = 1$), so the relevance matrix R_{SS} calculated by means of function g is different from the obtained one for Monday and Tuesday. Note that the NL expression $a_{SS_{22}}$ (*"the user has spent little time stopped"*) that has a validity degree of 1 and was relevant for Monday and Tuesday, now is irrelevant and is not included in the daily report. Here, the combination of W_{SS} and R_{SS} gives a different dynamic text report that contain the following information:

"On Wednesday, the user had a moderate lifestyle, however, he/she has spent little time walking. Nevertheless, he/she has spent quite a lot time in idle state, which could condition the results."

The rest of daily reports obtained from the same week are the following:

"On Thursday, the user had a vigorous lifestyle, because he/she has spent little time stopped and quite a lot time walking."
"On Friday, the user had an active lifestyle again, because he/she has spent little time stopped, a normal amount of time walking and little time running."
"On Saturday, the same as on Tuesday, the user had an active lifestyle, because he/she has also spent little time stopped and a normal amount of time walking. However, he/she has spent quite a lot time in idle state, which could condition the results."
"On Sunday, the user had a moderate lifestyle, because he/she has spent only very little time walking. However, he/she has spent too much time in idle state, which could condition the results."

As stated in Section 3.4, in medical applications the definition of the linguistic labels that summarize the amount of physical activity (PM_{SS}) have to be done by an expert according to the final goal. In order to show an example about the sensitivity of the choices in these daily reports, we have tuned the definition of linguistic labels that summarize *walking* state for two users completely different with respect to the experimentation one (user 1).

Firstly, we have figured out a user with reduced mobility (user 2) for whom short walks represent great efforts. Secondly, we have figured out the physical activity of an athlete (user 3) for whom daily walking for long distances is necessary. Table 6 shows the linguistic distribution of labels for these three users, jointly with the daily report obtained on Monday of the analyzed week.

Note that, as expected, for the same amount of physical activity the systems returns different reports, depending on the user's needs or goals. These differences concern not only the lifestyle results, but also the information content provided by the report.

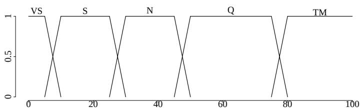
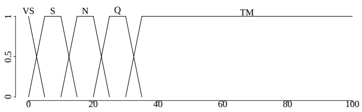
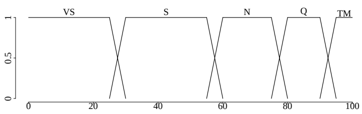
User	Walking definition	Daily report
1		*"On Monday, the user had an active lifestyle, because he/she has spent little time walking and little time running."*
2		*"On Monday, the user had a vigorous lifestyle, because he/she spent quite a lot time walking."*
3		*"On Monday, the user had a sedentary lifestyle, because he/she spent very little time walking and quite a lot time stopped."*

Table 6: Monday reports for the three analyzed different users.

5.2.3. Weekly reports

The dynamic character of weekly reports is determined by the relevance matrix R_{SDPA} instantiated by the g function of PM_{SDPA} (Table 3). The linguistic report obtained for the analyzed week is the following:

"In summary, this week the user had an active lifestyle, because some of the days the lifestyle was active and, on Wednesday and Sunday, the lifestyle decreased to moderate."

From the remaining 19 monitored weeks, we have selected two weekly reports that demonstrate the variability and the dynamic capacity of the developed tool to highlight the relevant information in each specific case:

"In summary, this week the user had a vigorous lifestyle, because on Monday he/she had an extreme lifestyle. However, during the weekend, the lifestyle decreased to moderate."

"In summary, this week the user had a moderate lifestyle, because all the days the lifestyle was moderate."

6. Conclusions

We have developed our research around the practical project of developing a new computational application able to self-track the physical activity of users and communicate the relevant information. The challenge of creating quality linguistic reports has led us to produce several contributions summarized as follows:

- We have designed a transparent linguistic model of the physical activity that provides a set of sentences with an associated validity and relevance degree.

- The developed model describes in everyday language the user's lifestyle, making reports interpretable by everyone, both experts and non-experts.
- We have extended our previous architecture by adding two new modules that calculate and highlight the relevant information according to the user's needs.
- The model is able to generate linguistic descriptions with different levels of temporal detail, adjusting the reports from the most concise to the most general, in each specific case.

There are some aspects of this work that could be addressed in a more detailed way in future works. One of these future research lines consists of conducting a formal analysis of the language used to create final reports. These reports must use a corpus of expressions that the user understands or commonly uses to describe the time series. On the other hand, in order to generate the best report for each specific user, the designing of a metric that objectively evaluate the quality of the reports is needed.

Despite this pending work, the performed experimentation demonstrates the potential and the flexibility of the application presented in this paper. In this paper we have developed a methodology able to produce dynamic linguistic descriptions of time series. The next step needs the collaboration of experts that help us to adjust some parameters and to guide the tool to specific application cases, like the physical monitoring of people with any type of physical or mental disorder. In this sense, this technique could be successfully applied to complement the emerging commercial products oriented to monitor the physical activity, providing a different and high quality way of analyzing and communicating results.

Acknowledgments

This work has been funded by the Spanish Government (MICINN) under project TIN2011-29827-C02-01 and the Principality of Asturias Government under the project CT13-52.

References

[1] W. Robertson, S. Stewart-Brown, E. Wilcock, M. Oldfield, M. Thorogood, Utility of accelerometers to measure physical activity in children attending an obesity treatment intervention, Journal of Obesity (2011).
[2] L. Bao, S. S. Intille, Activity recognition from user-annotated acceleration data, in: Proceedings of the 2nd International Conference on Pervasive Computing, pp. 1–17.
[3] K. Y. Chen, M. Sun, Improving energy expenditure estimation by using a triaxial accelerometer, Journal of applied Physiology 83 (1997) 2112–2122.
[4] J. Kwapisz, G. Weiss, S. Moore, Activity recognition using cell phone accelerometers, ACM SIGKDD Explorations Newsletter 12 (2011) 74–82.
[5] K. Ouchi, M. Doi, Living activity recognition using off-the-shelf sensors on mobile phones, Annals of Telecommunications (2012) 1–9.
[6] S.-B. Cho, K.-J. Kim, K. S. Hwang, I.-J. Song, Anidiary: Daily cartoon-style diary exploits bayesian networks, Pervasive Computing, IEEE 6 (2007) 66–75.
[7] H. Ketabdar, M. Lyra, System and methodology for using mobile phones in live remote monitoring of physical activities, in: IEEE International Symposium on Technology and Society (ISTAS), 2010, IEEE, pp. 350–356.
[8] N. Eagle, A. Pentland, Reality mining: sensing complex social systems, Personal and ubiquitous computing 10 (2006) 255–268.
[9] A. Mannini, A. M. Sabatini, Machine learning methods for classifying human physical activity from on-body accelerometers, Sensors 10 (2010) 1154–1175.
[10] N. Eagle, J. A. Quinn, A. Clauset, Methodologies for continuous cellular tower data analysis, in: Pervasive computing, Springer, 2009, pp. 342–353.
[11] K. Farrahi, D. Gatica-Perez, What did you do today?: discovering daily routines from large-scale mobile data, in: Proceedings of the 16th ACM international conference on Multimedia, ACM, pp. 849–852.
[12] M. Luštrek, B. Kaluža, Fall detection and activity recognition with machine learning, Informatica 33 (2009) 197–204.
[13] I. Anderson, H. Muller, Context awareness via gsm signal strength fluctuation, in: 4th international conference on pervasive computing, late breaking results, pp. 27–31.
[14] E. Miluzzo, N. D. Lane, K. Fodor, R. Peterson, H. Lu, M. Musolesi, S. B. Eisenman, X. Zheng, A. T. Campbell, Sensing meets mobile social networks: the design, implementation and evaluation of the cenceme application, in: Proceedings of the 6th ACM conference on Embedded network sensor systems, ACM, pp. 337–350.
[15] Nike Fuelband, http://www.nike.com/cdp/fuelband, 2013.

[16] Fitbit, http://www.fitbit.com, 2013.
[17] H. P. Grice, P. Cole, J. L. Morgan, Syntax and semantics, Logic and conversation 3 (1975) 41–58.
[18] M. Ros, M. Pegalajar, M. Delgado, A. Vila, D. Anderson, J. Keller, M. Popescu, Linguistic summarization of long-term trends for understanding change in human behavior, in: 2011 IEEE International Conference on Fuzzy Systems, pp. 2080–2087.
[19] D. Wu, J. Mendel, Linguistic summarization using if-then rules and interval type-2 fuzzy sets, Fuzzy Systems, IEEE Transactions on 19 (2011) 136–151.
[20] J. Kacprzyk, A. Wilbik, S. Zadrozny, Linguistic summarization of time series using a fuzzy quantifier driven aggregation, Fuzzy Sets and Systems 159 (2008) 1485–1499.
[21] A. Wilbik, J. M. Keller, G. L. Alexander, Linguistic summarization of sensor data for eldercare, in: Systems, Man, and Cybernetics (SMC), 2011 IEEE International Conference on, IEEE, pp. 2595–2599.
[22] R. Castillo-Ortega, N. Marín, D. Sánchez, A fuzzy approach to the linguistic summarization of time series., Multiple-Valued Logic and Soft Computing 17 (2011) 157–182.
[23] L. Eciolaza, M. Pereira-Farina, G. Trivino, Automatic linguistic reporting in driving simulation environments, Applied Soft Computing (2012).
[24] A. Alvarez-Alvarez, D. Sanchez-Valdes, G. Trivino, A. Sánchez, P. D. Suárez, Automatic linguistic report about the traffic evolution in roads, Expert Systems with Applications 39 (2012) 11293–11302.
[25] D. Sanchez-Valdes, A. Alvarez-Alvarez, G. Trivino, Linguistic description about circular structures of the Mars' surface, Applied Soft Computing 13 (2013) 4738–4749.
[26] L. Arguelles, G. Trivino, I-struve: Automatic linguistic descriptions of visual double stars, Engineering Applications of Artificial Intelligence 26 (2013) 2083–2092.
[27] A. Alvarez-Alvarez, G. Trivino, Linguistic description of the human gait quality, Engineering Applications of Artificial Intelligence 26 (2013) 13–23.
[28] G. Bailador, G. Trivino, Pattern recognition using temporal fuzzy automata, Fuzzy Sets and Systems 161 (2010) 37–55.
[29] D. Sanchez-Valdes, L. Eciolaza, G. Trivino, Linguistic description of human activity based on mobile phone's accelerometers, in: Ambient Assisted Living and Home Care, Springer, 2012, pp. 346–353.
[30] A. Alvarez-Alvarez, J. M. Alonso, G. Trivino, Human activity recognition in indoor environments by means of fusing information extracted from intensity of wifi signal and accelerations, Information Sciences 233 (2012) 162–182.
[31] L. A. Zadeh, From computing with numbers to computing with words - from manipulation of measurements to manipulation of perceptions, IEEE Transactions on Circuits and Systems 45(1) (1999).
[32] L. A. Zadeh, A new direction in ai. towards a computational theory of perceptios of measurements to manipulation of perceptions, AI Magazine 22(1) (2001).
[33] L. A. Zadeh, Toward human level machine intelligence - is it achievable? the need for a paradigm shift, IEEE Computational Intelligence Magazine (2008).
[34] L. A. Zadeh, The concept of a linguistic variable and its application to approximate reasoning, Parts I, II, and III. Information Sciences 8, 8, 9 (1975) 199–249, 301–357, 43–80.
[35] FAO/WHO-OMS/UNU, Expert consultation report. Energy and Protein Requirements, Technical Report Series 724 (1985).
[36] D. Sanchez-Valdes, G. Trivino, Computational perceptions of uninterpretable data. a case study on the linguistic modeling of human gait as a quasi-periodic phenomenon, Fuzzy Sets and Systems 253 (2014) 101 – 121.
[37] A. Alvarez-Alvarez, G. Trivino, O. Cordón, Human gait modeling using a genetic fuzzy finite state machine, Fuzzy Systems, IEEE Transactions on 20 (2012) 205–223.
[38] C. Alsina, M. J. Frank, B. Schweizer, Associative Functions: Triangular Norms and Copulas, World Scientific Publishing Company, Singapore, 2006.
[39] M. Delgado, D. Sánchez, M. Vila, Fuzzy cardinality based evaluation of quantified sentences, Int. Journal of Approximate Reasoning 23 (2000) 23–66.
[40] L. A. Zadeh, A computational approach to fuzzy quantifiers in natural languages, Computing and Mathematics with Applications 9 (1983) 149–184.
[41] E. Reiter, R. Dale, Z. Feng, Building natural language generation systems, MIT Press, 2000.

7.4. Linguistic and emotional feedback for self-tracking physical activity

D. Sanchez-Valdes, A. Alvarez-Alvarez, and G. Trivino. "Linguistic and emotional feedback for self-tracking physical activity". Submitted to *Expert Systems with Applications*, 2015.

Linguistic and emotional feedback for self-tracking physical activity

Daniel Sanchez-Valdes[a,*], Gracian Trivino[a]

[a] *European Centre for Soft Computing,*
Edificio de Investigación, c/ Gonzalo Gutiérrez Quirós s/n
33600 Mieres, Asturias, Spain (+34 985 45 65 45).

Abstract

The level of physical activity is one of the most important parameters to monitor the evolution of some physical and mental disorders, e.g., obesity or depression. Its continuous monitoring provides an objective and quantitative measurement tool to be applied by therapists when developing patient-specific treatments. In addition, the use of virtual agents that interact and communicate relevant information in an emotional way, increases the effectiveness and acceptance of the Human Computer Interfaces. Here, we develop a computational application that combines activity recognition techniques with affective computing ones. We include first experimental results that demonstrate the potential and applicability of this tool.

Keywords: Self-tracking, Activity Recognition, Affective Computing, Linguistic Description of Data

1. Introduction

The concept of self-tracking, quantified self or selflogging is emerging in recent years thanks to the incorporation of technology advances in data acquisition. It allows users to acquire a better self-knowledge, as well as it allows experts (either endocrinologists, psychologists, personal trainers, or others) to adjust treatments and analyze the physical evolution of their patients, for example, when analyzing results of medical interventions.

More specifically, activity recognition has captured the attention of Computer Science communities in connection to different fields of study such as medicine, human-computer interaction, or sociology. It can help to address health concerns that arise due to inactivity, e.g., cardiovascular diseases, hypertension, osteoporosis and the critical public health threat of childhood obesity (Robertson et al., 2011). The majority of the population fails to meet the recommended levels for good health habits. For this purpose, the design of practical and reliable methods to monitor and motivate individual's daily physical activities could be very useful. Following subsections present the two main components of this challenge: first, how to measure physical activity and, then, how to communicate results in the most effective way.

1.1. Measuring physical activity and estimating energy requirements

We define the level of physical activity as a measure dependent on the duration, frequency and intensity of the physical movements performed by a person during a period of time, typically one day.

The best objective approach to assess the level of daily physical activity is based on registering people movements with the help of sensors, including measures from heart rate monitors, pedometers, gyroscopes, magnetometers and accelerometers. There are also subjective methods that consist of recalling questionnaires like the International Physical Activity Questionnaire and physical activity logbooks (Booth et al., 2003).

*Corresponding author
Email addresses: `daniel.sanchezv@softcomputing.es` (Daniel Sanchez-Valdes), `gracian.trivino@softcomputing.es` (Gracian Trivino)

Preprint submitted to Expert Systems with Applications *November 28, 2014*

Most of existing activity recognition systems have been developed using multiple accelerometers strapped to the subject's extremities (Bao and Intille, 2004). By this way, a good number of activities are recognizable but it is impracticable for users to continuously wear so many accelerometers over their bodies in daily life. Nowadays, smartphones and other electronic devices contain diverse and powerful sensors, opening up exciting areas for research in data mining and human-computer interaction. Kwapisz et al. (2011) describe and evaluate a system that uses the accelerations taken from smartphones situated in the user's trousers pocket. With these data, they could label daily activities such as walking, running, climbing stairs, sitting and standing. Ouchi and Doi (2012) performed a living activity recognition using accelerometers and microphones of smartphones, to determine whether the user was walking, quiet or performing a task by wearing the mobile device in the breast pocket. Kwon et al. (2014) proposed an approach based on unsupervised learning methods for human activity recognition using a smartphone kept in a pants pocket.

We have observed two important limitations of these currently available solutions. First, all of these solutions have a common limitation related to the location of sensors over the body. All of them require a specific location of the device, such as the trousers pocket, the breast pocket or the subject's waist, limiting the functionality of the system. The development of algorithms that do not require a specific location of the device is needed from the functional point of view. Second, most smartphone based activity recognition systems apply statistical machine learning models to identify and label the activities, like in Unold et al. (2011), where authors used a machine learning approach combining Fuzzy Logic with an immune algorithm to model sport training. Most machine learning models infer results without providing any knowledge or explanation of they internal working (black box systems). The development of a system with a “transparent” internal working could be really interesting from the interpretability point of view.

1.2. Communicating results in an emotional way

Once we know which is the level of physical activity, we have to communicate it to the user in the most appropriate way. Most of existing solutions provide statistical information, e.g., numerical data, graphics and tables, but a new solution that linguistically describes the information is needed for either therapists or patients following a therapy. To deal with this necessity, Linguistic Description of Data is intended, in general, for applications in which there is a strong human-computer interaction involving accessing and understanding data (Reiter and Dale, 2000).

Moreover, when we desire to assist people in modifying their health behavior and their physical routines, a suitable motivation is essential. Thus, Fogg (2002) describes the idea of “computer based tools designed for the purpose of changing people's attitudes and behaviors”. Technology can encourage changes in physical activity through meaningful and motivational messages, attractive visuals or credits in gaming systems. The way to motivate people to reach their objectives varies depending on the specific goals, the application domain and the user's physical and mental conditions. In addition, recommendations must take into account the cultural, social and environmental characteristics of each user. Applications that need an appropriate motivation could range from those that promote the loss of weight in people with obesity problems to those that promote the maintaining of exercise routines in people with depression.

To deal with this challenge, the use of emotions plays a significant role in the human decision-making process as an important aspect of the human intelligence (El-Nasr et al., 2000; Damasio, 2008). Sensitivity and expressiveness are very important in human interaction and transmit valuable information (Picard, 2000). The best known emotion synthesis technologies usually trigger visible or verbal displays of emotions. In this way, a variety of emotion models have been proposed in areas such as Cognitive Science, Philosophy and Artificial Intelligence.

One of the main goals in Artificial Intelligence is to build intelligent systems that act and reason as humans in a specific domain. Virtual agents are used in many applications that require human-like interfaces and offer a technology that dynamically interacts with people to gather information, monitor health care, provide information, or even act as companions (Martınez-Miranda and Aldea, 2005; Alepis and Virvou, 2013). Sensors are used to monitor the environment and virtual systems communicate with users in a manner that has a powerful effect on their living situation. In Kenny et al. (2009) and Mahamood and Reiter (2011) we find two research examples focused on virtual agents able to produce emotional utterances.

This work is a first approach of a long project in the way of motivating changes in users' physical behavior. Here, we design a computational model of emotions that takes as input the physical activity of users and communicate results in an emotional way, that is, by changing mood, linguistic and facial expressions in order to improve the empathy with the user. The aim of this work is, therefore, design a virtual agent whose emotions vary depending on the users' physical activity, which is in itself a contribution in the field. As future research, we will analyze the effectiveness that these emotional messages have in users physical behavior, and how they can help users to reach the objectives established by an expert.

Within the design of computational systems that linguistically describe phenomena, our research line is based on the Computational Theory of Perceptions (CTP) introduced by Zadeh (1999) in the seminal paper "From computing with numbers to computing with words - from manipulation of measurements to manipulation of perceptions" and further developed in subsequent papers (Zadeh, 2008). In previous research, we have worked with accelerometer data by automatically generating linguistic reports about human gait quality (Alvarez-Alvarez and Trivino, 2013), gesture recognition (Bailador and Trivino, 2010) and activity recognition (Alvarez-Alvarez et al., 2013), among other application fields. Attending to affective computing, we have simulated emotional personality in Human Computer Interfaces (van der Heide and Trivino, 2010).

1.3. Contributions of this work

The main contributions of this paper are summarized as follows:

- We have used current smartphone's accelerometers to acquire and store the accelerations produced during the daily activity of a person, without any limitation of location. We have designed an interpretable and transparent model of human activity that describes patterns emerging in data and estimates the level of physical activity of a person and his/her energy requirements. Section 3 explains this model as an alternative to current black box activity recognition approaches.
- We have extended the architecture used in previous works on linguistic description of complex phenomena to incorporate an emotional response to the user's behavior and necessities. Section 4 describes the virtual agent, or avatar, whose emotional state changes depending on input data. This avatar complements facial gestures by adding advices with emotional content that the user receives for self-tracking his/her physical activity.
- We have developed a first practical experimentation that shows the feasibility of our approach. Section 6 presents our prototype and the experimental results.

2. General architecture

Fig. 1 shows our architecture for generating linguistic descriptions of phenomena that has been instantiated to the goals of this paper. In this architecture, processes are graphically represented with ovals and data structures are represented with rectangles. The two main stages in the report generation are explained as follows:

2.1. Off-line building process

Our approach is based on the subjective perceptions of a domain expert that we call the designer. The more experienced designer, with better understanding and use of Natural Language in the application domain, the richer the model with more possibilities of achieving and responding to final users' needs and expectations. Here, the designer is assisted by therapists, psychologists and linguistics experts.

During the *Physical activity analysis phase* the designer collects a corpus of Natural Language expressions that are typically used to describe the relevant features of the monitored phenomenon. The analysis of the particular meaning of each linguistic expression allows to build a Granular Linguistic Model of the Phenomenon (GLMP), that here, is related to the physical activity of the user ($GLMP_{Activity}$). The GLMP data structure is a useful paradigm that contains a collection of valid sentences to describe the phenomenon with different granularity degrees (Alvarez-Alvarez et al., 2012; Sanchez-Valdes et al., 2013).

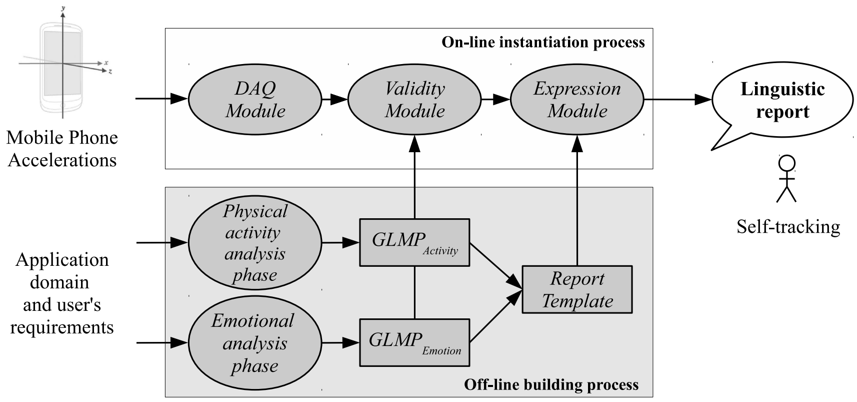


Figure 1: Architecture of computational systems for generating linguistic description of human activity based on acceleration data.

During the *Emotional analysis phase*, the user's requirements help the designer to collect a corpus of expressions with emotional content able to motivate responses on the user's behavior. These linguistic expressions allow building the Granular Linguistic Model of an emotional agent ($GLMP_{Emotion}$).

The designer builds a *Report Template* that allows combining the meaning of the valid sentences in the $GLMP_{Activity}$ and the $GLMP_{Emotion}$ into relevant descriptions. These reports must be tuned in order to cover the necessities and use of language of each different therapy and type of user.

2.2. On-line instantiation process

The *Data Acquisition (DAQ) Module* obtains numerical data from the triaxial accelerometers embedded in current smartphones. We have developed an experimental application based in Android platform that acquires and stores the accelerations a_x, a_y and a_z, in m/s^2, produced during the daily routine activities of a person, e.g., walking, standing or running. The use of smartphones allows exploiting a great range of communications capabilities, integrated hardware and software features.

The DAQ Module calculates the accelerations module (ρ) and the standard deviation of data (σ) through an experimentally obtained 20 seconds moving window. It is worth remarking that ρ and σ are invariant with respect to the smartphone orientation and position, allowing estimating the level of physical activity by wearing the smartphone wherever the user wants.

The *Validity Module* takes the numerical inputs acquired by the *DAQ Module* and creates an instance of the $GLMP_{Activity}$ and the $GLMP_{Emotion}$. The module provides a set of linguistic expressions that describe the current state of the phenomenon with an associated validity degree.

The *Expression Module*, according to the *Report Template*, takes this set of linguistic expressions, selects and organizes the most relevant ones, and generates a linguistic report that helps the user to self-track his/her physical activity.

3. Granular Linguistic Model of the physical activity

The main element of this structure is known as Computational Perception (CP), which is based on the concept of linguistic variable developed by Zadeh (1975). CPs are computational models of units of information (granules) acquired by the designer about the phenomenon to be modeled, i.e., CPs correspond to perceptions of specific parts of the phenomenon at certain granularity degree. A CP is a tuple (A, W) described as follows:

$A = (a_0, a_1, a_2, \ldots, a_n)$ is a vector of linguistic expressions (words or sentences in Natural Language) that represent the whole linguistic domain of CP. During the off-line building process, the components of A are defined by the designer in accordance to the most suitable sentences from the typically used ones in the application domain of language.

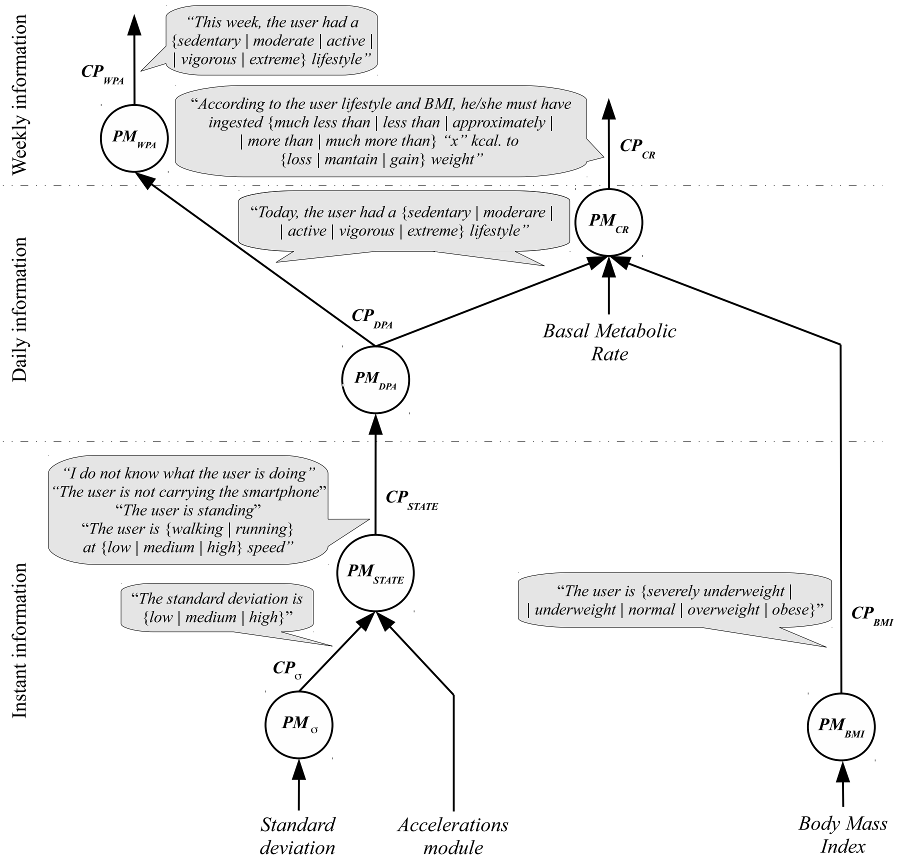


Figure 2: $GLMP_{Activity}$ that linguistically describes the physical activity.

$W = (w_0, w_1, w_2, \ldots, w_n)$ is a vector of validity degrees $w_i \in [0, 1]$ assigned to each a_i during the on-line instantiation process. In the application context, w_i represents the suitability of a_i to describe the current perception. Since the designer chooses the components of A by trying to cover all the possible values of CP, the total validity should be distributed among all linguistic labels. Therefore, typically, the components of A are attached to fuzzy sets forming strong fuzzy partitions (Ruspini, 1969) in the universe of discourse of CP, i.e., $\sum_{i=0}^{n} w_i = 1$.

The GLMP consists of a network of Perception Mappings (PM), which are the elements used to create and aggregate CPs. Here, the $GLMP_{Activity}$ (Fig. 2) collects the set of concepts and perceptions needed to define the physical activity of users. In this application, we deal with monitoring the level of physical activity of a person by analyzing his/her basic movements, such as, walking, running and standing. The model also identifies when the smartphone used to capture data is not with the user, e.g., when the device has been left on a desk. The $GLMP_{Activity}$ provides three different granularity levels: instant, daily or weekly information. The following subsections detail each PM included in the $GLMP_{Activity}$:

3.1. Standard deviation of accelerations module (PM_σ)

In the case of this PM_σ, the inputs are numerical data provided by the DAQ Module. PM_σ is a tuple (U, y, g, T), where each component is explained as follows:

U is a time series of numerical input data corresponding to the standard deviation (σ) of the accelerations module.

y is the output $CP_\sigma = (A_y, W_y)$, where $A_y =$(*Low (L), Medium (M), High (H)*).

g is the output function $W_y = g(\sigma)$ obtained by means of trapezoidal MFs forming a strong fuzzy partition. The linguistic labels of A_y are defined by their vertices as follows: $\{L\ (0, 0, 0.001, 0.015)$, M $(0.001, 0.015, 0.08, 0.095)$, $H\ (0.08, 0.095, \infty, \infty)\}$.

T is a text generation algorithm that allows generating the sentences in A_y. Here, T produces linguistic expressions as follows: *"The standard deviation is {low | medium | high}"*.

3.2. Body Mass Index (PM_{BMI})

U is the Body Mass Index (BMI) of the user.

y is the output CP_{BMI}, where A_y =(*Severely underweight (SUW), Underweight (UW), Normal (N), Overweight (OW), Obese (O)*).

g is the output function that calculates the validity degrees of the output CP_{BMI}. The linguistic labels of A_y and their definition are established according to the BMI International Classification provided by the World Health Organization (WHO, 2004) as follows: $\{SUW\ (-\infty, -\infty, 15.5, 16.5)$, UW $(15.5, 16.5, 18, 19)$, $N\ (18, 19, 24.5, 25.5)$, $OW\ (24.5, 25.5, 29.5, 30.5)$, $O\ (29.5, 30.5, \infty, \infty)\}$.

T produces linguistic expressions as follows: *"The user is {severely underweight | underweight | normal | overweight | obese}"*.

3.3. Physical states (PM_{STATE})

In this PM_{STATE} we identify three of the basic physical activities established in FAO/WHO-OMS/UNU (1985): *Standing*, *Walking* and *Running*. In addition, we define other two possible states, namely, the *Unknown* and the *Idle* states. The *Unknown* state deals with representing the non-interpretability of the signal, providing high robustness to the system (for more information about uninterpretable data see Sanchez-Valdes and Trivino (2013)). It also represents those situations in which data is not captured, e.g., when the smartphone's sensor is malfunctioning or the device is switched off. The *Idle* state represents those instants in which the user is not carrying the smartphone, e.g., when he/she leaves it on a table. This PM has the following elements:

U is a vector of input CPs, $U = (u_1, u_2, \ldots, u_m)$, where u_i are tuples (A_i, W_i) and m the number of input CPs. Here, U is the input CP_σ and the numerical input data corresponding to the accelerations module (ρ).

y is the output CP_{STATE}, where A_y =(*Unknown* (q_0), *Idle* (q_1), *Standing* (q_2), *Walking* (q_3), *Running* (q_4)).

g is the aggregation function $W_y = g(W_1, W_2, \ldots, W_m)$, where W_i are the vectors of validity degrees of the m input CPs. Here, we use a Fuzzy Finite State Machine (FFSM) to implement this aggregation function. For a more detailed description for the paradigm of FFSM and its applications, the interested reader could see the paper (Alvarez-Alvarez et al., 2012).

According to the states diagram shown in Fig. 3, g is composed by rules R_{ii} to remain in the state q_i and rules R_{ij} to change from the state q_i to the state q_j. These fuzzy rules have the following structure:

$$\begin{aligned} R_{ii}: &\ \text{IF } (\mathit{State}[t-1] \text{ is } q_i) \text{ AND } (\mathit{Input\ variables\ constraints})_{ii} \text{ AND } (\mathit{Temporal\ constraints})_{ii} \\ &\ \text{THEN } (\mathit{State}[t] \text{ is } q_i) \\ R_{ij}: &\ \text{IF } (\mathit{State}[t-1] \text{ is } q_i) \text{ AND } (\mathit{Input\ variables\ constraints})_{ij} \text{ AND } (\mathit{Temporal\ constraints})_{ij} \\ &\ \text{THEN } (\mathit{State}[t] \text{ is } q_j) \end{aligned} \tag{1}$$

The antecedent of each rule is composed by the state of the signal at timestamp $t-1$, the input variables constraints, and the temporal constraints, which are described below in Section 3.3.1 and Section 3.3.2, respectively. In Section 3.3.3, we detail the set of fuzzy rules.

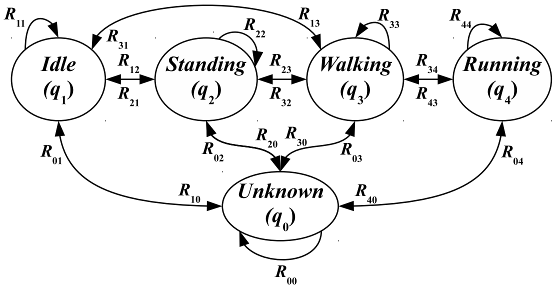


Figure 3: States diagram of the FFSM.

T produces linguistic expressions that can be adapted depending on the situation. When the model is not able to recognize the input signal (q_0), the system reports the following expression: *"We do not know what the user is doing"*. However, when the system detects that the user does not carry the smartphone (q_1), the report is: *"The user is not carrying the smartphone"*. When the signal is in the state q_2, the template is: *"The user is standing"*. Finally, when the physical activity corresponds to the states q_3 or q_4, the template is: *"The user is {walking | running} at {low | medium | high} speed."*.

3.3.1. Constraints on the input variables

The input variable constraints are defined over the standard deviation of the accelerations module (CP_σ) and the cadence when the user is walking or running. This cadence is obtained by applying the Fast Fourier Transform (FFT) to the accelerations module during a temporal moving window when CP_σ is *high*.

We consider that the user is walking when the cadence is defined as follows: {*extremely low* $(0, 0, 65, 75)$, *very low* $(65, 75, 105, 115)$, *low* $(105, 115, 130, 135)$}. We consider that the user is running when the is defined as follows: {*high* $(130, 135, 140, 145)$, *very high* $(140, 145, 160, 165)$, *extremely high* $(160, 165, \infty, \infty)$}.

3.3.2. Temporal constraints

To provide robustness to the system, we have introduced temporal constraints over the length of the states. We numerically compute the duration d_{q_j} as the time that the constraints on input variables of q_j are fulfilled, increasing d_{q_j} according to the sampling period.

Temporal fuzzy constraints represent the minimum duration (in seconds) that the input variables constraints have to be stable to consider the change from state q_i to state q_j. Thus, we have defined the linguistic labels *enough* and *not enough* to indicate when the duration d_{q_j} of each state q_j is enough or not to consider the change to this state. Each possible change of state has its own definition of these two linguistic labels: d_{q_1} is {*enough* $(-\infty, -\infty, 290, 310)$, *not enough* $(290, 310, \infty, \infty)$}, d_{q_2} is {*enough* $(-\infty, -\infty, 20, 40)$, *not enough* $(20, 40, \infty, \infty)$}, d_{q_3} is {*enough* $(-\infty, -\infty, 10, 15)$, *not enough* $(10, 15, \infty, \infty)$}, and d_{q_4} is {*enough* $(-\infty, -\infty, 10, 15)$, *not enough* $(10, 15, \infty, \infty)$}.

3.3.3. Set of fuzzy rules

According to the structure presented in (1), the rule base that models the physical activity is the following:

R_{01}: IF ($S[t]$ is *unknown*) AND (σ is *low*) AND (d_{q_1} is *enough*) THEN ($S[t+1]$ is *idle*)

R_{11}: IF ($S[t]$ is *idle*) AND $\Big($(σ is *low*) OR (σ is *medium* AND d_{q_2} is *not enough*) OR (σ is *high* AND d_{q_3} is *not enough*)$\Big)$ THEN ($S[t+1]$ is *idle*)

R_{21}: IF ($S[t]$ is *standing*) AND (σ is *low*) AND (d_{q_1} is *enough*) THEN ($S[t+1]$ is *idle*)

R_{31}: IF ($S[t]$ is *walking*) AND (σ is *low*) AND (d_{q_1} is *enough*) THEN ($S[t+1]$ is *idle*)

R_{02}: IF ($S[t]$ is *unknown*) AND (σ is *medium*) AND (d_{q_2} is *enough*) THEN ($S[t+1]$ is *standing*)

R_{12}: IF ($S[t]$ is *idle*) AND (σ is *medium*) AND (d_{q_2} is *enough*) THEN ($S[t+1]$ is *standing*)

R_{22}: IF ($S[t]$ is *standing*) AND $\Big($(σ is *low* AND d_{q_1} is *not enough*) OR (σ is *medium*) OR (σ is *high* AND d_{q_3} is *not enough*)$\Big)$ THEN ($S[t+1]$ is *standing*)

R_{32}: IF ($S[t]$ is *walking*) AND (σ is *medium*) AND (d_{q_2} is *enough*) THEN ($S[t+1]$ is *standing*)

R_{03}: IF ($S[t]$ is *unknown*) AND (σ is *high* AND FFT_ρ is $\{$*extremely low* | *very low* | *low*$\}$ AND d_{q_3} is *enough*) THEN ($S[t+1]$ is *walking at* $\{$*low* | *medium* | *high*$\}$ *speed*)

R_{13}: IF ($S[t]$ is *idle*) AND (σ is *high* AND FFT_ρ is $\{$*extremely low* | *very low* | *low*$\}$ AND d_{q_3} is *enough*) THEN ($S[t+1]$ is *walking at* $\{$*low* | *medium* | *high*$\}$ *speed*)

R_{23}: IF ($S[t]$ is *standing*) AND (σ is *high* AND FFT_ρ is $\{$*extremely low* | *very low* | *low*$\}$ AND d_{q_3} is *enough*) THEN ($S[t+1]$ is *walking at* $\{$*low* | *medium* | *high*$\}$ *speed*)

R_{33}: IF ($S[t]$ is *walking*) AND $\Big($(σ is *low* AND d_{q_1} is *not enough*) OR (σ is *medium* AND d_{q_2} is *not enough*) OR (σ is *high* AND FFT_ρ is $\{$*extremely low* | *very low* | *low*$\}$)$\Big)$ THEN ($S[t+1]$ is *walking at* $\{$*low* | *medium* | *high*$\}$ *speed*)

R_{43}: IF ($S[t]$ is *running*) AND (σ is *high* AND FFT_ρ is $\{$*extremely low* | *very low* | *low*$\}$ AND d_{q_3} is *enough*) THEN ($S[t+1]$ is *walking at* $\{$*low* | *medium* | *high speed*$\}$)

R_{04}: IF ($S[t]$ is *unknown*) AND (σ is *high* AND FFT_ρ is $\{$*high* | *very high* | *extremely high*$\}$ AND $d_{Running}$ is *enough*) THEN ($S[t+1]$ is *running at* $\{$*low* | *medium* | *high speed*$\}$)

R_{34}: IF ($S[t]$ is *walking*) AND (σ is *high* AND FFT_ρ is $\{$*high* | *very high* | *extremely high*$\}$ AND $d_{Running}$ is *enough*) THEN ($S[t+1]$ is *running at* $\{$*low* | *medium* | *high speed*$\}$)

R_{44}: IF ($S[t]$ is *running*) AND (σ is *high* AND FFT_ρ is $\{$*high* | *very high* | *extremely high*$\}$ AND $d_{Running}$ is *enough*) THEN ($S[t+1]$ is *running at* $\{$*low* | *medium* | *high speed*$\}$)

R_{i0}: ELSE ($S[t+1]$ is *unknown*)

The final values of $S[t+1]$ are calculated as a weighted average of the individual rules, where the weight of each rule R_{ij} corresponds to its firing degree τ_{ij}. This firing degree is calculated using the minimum for the AND operator and the bounded sum of Łukasiewicz (Alsina et al., 2006) for the OR operator.

3.4. Daily physical activity (PM_{DPA})

Once the model recognizes the physical activities, this PM estimates the user's level of daily physical activity (DPA). It has the following elements:

U is the input CP_{STATE}.

y is the output CP_{DPA}, where A_y =(*Sedentary (S), Moderate (M), Active (A), Vigorous (V), Extreme (E)*).

g is the output function that estimates the user's daily physical activity by means of the Eq. 2:

$$\text{DPA} = \frac{\sum_{i=0}^{4} TD_i \cdot UW \cdot AF \cdot GF}{t} \tag{2}$$

where i is the recognized activity (0=*Unknown*, 1=*Idle*, 2=*Standing*, 3=*Walking* and 4=*Running*), TD_i is the total duration that the user has practiced the activity i during the recording period t (in minutes), UW is the user's weight expressed in kilograms and GF is a gender factor that takes value 1 if the user is a man and 0.9 if she is a woman. AF is a factor whose value depends on the activity as follows: 0.0285 (*Unknown*, *Idle* and *Standing*), 0.038 (*Walk slowly*), 0.063 (*Walk normal*), 0.106 (*Walk fast*), 0.15 (*Run slowly*), 0.2 (*Run normal*) and 0.25 (*Run fast*). During the *Unknown* or *Idle* states, we consider that the level of physical activity is, at least, equal to the minimal level of physical activity, that it is produced when he/she is *Standing*. This way to estimate the physical activity has been obtained from FAO/WHO-OMS/UNU (1985), where a similar but crisp way is proposed to estimate the energy consumption.

As seen in (2), the levels of physical activity vary depending on the gender and weight of the user. In consequence, the distribution of the linguistic labels that define the user's lifestyle has to be tuned according to the user's gender and weight. Their vertices are the following: $\{S\ (0, 0, 2.23C_1, 2.37C_1)$, M $(2.23C_1, 2.37C_1, 2.58C_1, 2.72C_1)$, $A\ (2.58C_1, 2.72C_1, 2.93C_1, 3.07C_1)$, $V\ (2.93C_1, 3.07C_1, 3.28C_1, 3.42C_1)$, $E\ (3.28C_1, 3.42C_1, \infty, \infty)\}$, where C_1 is a scale factor calculated as follows:

$$C_1 = \frac{UW \cdot GF}{80} \tag{3}$$

UW is the user's weight (in kilograms), GF is the gender factor explained above, and 80 is the weight of the user taken as reference to define the linguistic labels.

T produces linguistic expressions as follows: *"Today, the user had a {sedentary | moderate | active | vigorous | extreme} lifestyle"*.

3.5. Calorie requirements (PM_{CR})

This PM allows us to inform about the amount of calories that the user needs to intake in order to maintain his/her weight according to the estimated level of physical activity. This PM_{CR} has the following elements:

U are the input CPs: $\{CP_{BMI}, CP_{DPA}\}$ and the vector of numerical input data corresponding to the Basal Metabolic Rate (BMR) of the user.

y is the output CP_{EA}, where A_y =(*Much less (ML), Less (L), Approximately (A), More (M), Much more (MM)*).

g is the output function that calculates the validity degrees of the output CP_{CR}. Here, they are the same as the defined in PM_{BMI}.

The Calorie requirements (CR) are estimated by combining the BMR with the level of physical activity, as follows:

$$\text{CR (kcal)} = \text{BMR} \times \sum_{i=1}^{5} PAF \cdot w_{DPA_i} \tag{4}$$

where i represents each lifestyle modeled in PM_{DPA} (1=*Sedentary*, 2=*Moderate*, 3=*Active*, 4=*Vigorous* and 5=*Extreme*) and PAF is the physical activity factor that corresponds to each of them (1.2, 1.375, 1.55, 1.725 and 1.9, respectively). This way of estimating the CR is taken from Mifflin et al. (1990), where we have adapted the crisp approach to the fuzzy modeling.

Then, we contrast this information with PM_{BMI}. If the user is *underweight*, or *severely underweight*, he/she needs to increase weight, so the calories intake must be higher, or much higher, than the calculated CR. On the other hand, if the user has a *normal* weight, he/she must intake approximately the amount of calories calculated to maintain weight. Finally, if the user is *overweight*, or *obese*, he/she needs to decrease weight, so the calories intake must be lower, or much lower, than the calculated CR.

T produces linguistic expressions as follows: *"According to the user lifestyle and BMI, he/she must have ingested {much less than | less than | approximately | more than | much more than} "x" kcal. to {loss | maintain | gain} weight"*, being "*x*" the number of calories calculated in CR. It is worth remarking that, these values have to be adjusted by the therapist according to the specific user's requirements.

3.6. Weekly physical activity (PM_{WPA})

U is a vector of the numerical values of DPA.

y is the output CP_{WPA}, where A_y =(*Sedentary (S), Moderate (M), Active (A), Vigorous (V), Extreme (E)*).

g is the output function that calculates the validity degrees of the output CP_{WPA}. We aggregate the DPA corresponding to each day of the week by means of the weighted average represented by Eq. 5. Here, since we are interested in highlighting extreme behaviors like sedentary or hyperactive lifestyles, we have weighted them with higher coefficients. These coefficients can be tuned by the therapist in order to highlight other types of lifestyles according to each user pathology.

$$\text{WPA} = \frac{\sum_{i=1}^{7} K_i \cdot DPA_i}{\sum_{i=1}^{7} K_i} \tag{5}$$

$$K_i = 10 \cdot (w_{DPA_1})_i + 5 \cdot (w_{DPA_2})_i + (w_{DPA_3})_i + 5 \cdot (w_{DPA_4})_i + 10 \cdot (w_{DPA_5})_i$$

where, DPA_i is the DPA corresponding to each day i and $(w_{DPA_1})_i, (w_{DPA_2})_i, \ldots$ are the validity degrees of the output CP_{DPA} calculated in Section 3.4. The distribution of the linguistic labels is exactly the same of the used in that Section 3.4.

T produces linguistic expressions as follows: *"This week, the user had a* {*sedentary* | *moderate* | *active* | *vigorous* | *extreme*} *lifestyle"*.

4. Granular Linguistic Model of the emotional agent

There is a general agreement about the universality of emotions (Fragopanagos and Taylor, 2005). Some models such as the Plutchik Wheel (Plutchik, 1980), the Russell's circumplex model (Russell, 1979) and the Whissell's one (Whissell, 1989), are quite far of being perfect models but, according with the bibliography, provide a good balance between complexity and functionality. Here, based on these models, we have designed a simple one, where a set of selected emotions are situated in a circle around the center point, which represents a neutral emotion-free state (Fig. 4).

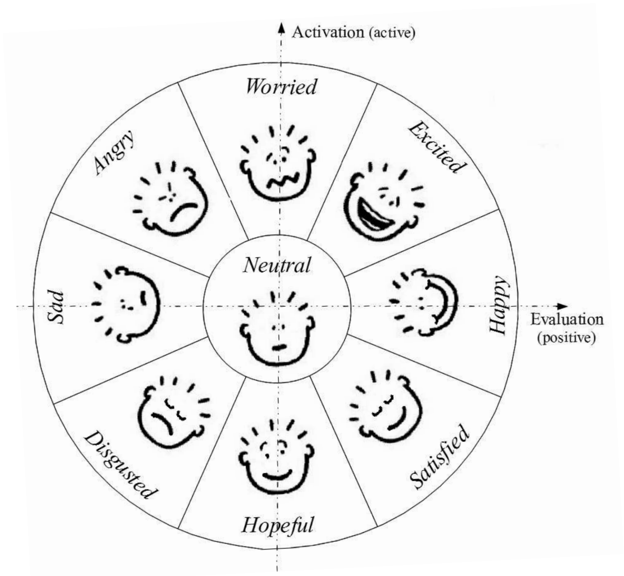


Figure 4: Emotional states and their facial expression.

The simulated emotional state of the avatar is categorized along two axes, namely, evaluation and activation. The evaluation axis, also known as valence, goes from negative to positive and represents the evaluation of the obtained results according to the specific requirements established by the therapist. The activation axis, also known as arousal, goes from passive to active and represents the strength of the avatar when communicating the evaluation results according to the therapist intentions.

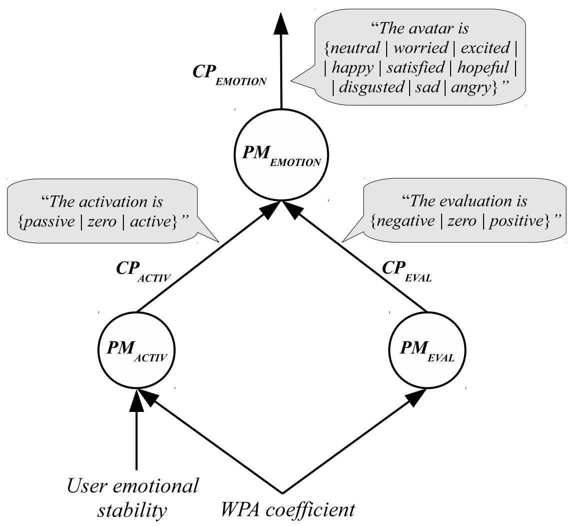


Figure 5: $GLMP_{Emotion}$ that models the avatar's emotional state.

Fuzzy Logic is an appropriate tool for modeling emotions since it provides an expressive language that enables our model to produce complex emotional states and behaviors (El-Nasr et al., 2000). In addition, the emotional state of a virtual agent is fuzzy, i.e., it can be in various emotional states at any instant of time.

Fig. 5 shows the $GLMP_{Emotion}$ that using the output of the $GLMP_{Activity}$, models the emotional state of a virtual agent and provides advices and recommendations to the user according to the therapist intentions, promoting the self-tracking of the user's physical activity. The following subsections detail each PM included in this $GLMP_{Emotion}$:

4.1. Evaluation of weekly results (PM_{EVAL})

This PM provides the evaluation of the obtained results. It has the following elements:

U is the numerical value of WPA.

y is the output CP_{EVAL}, where A_y =(*Negative (N), Zero (Z), Positive (P)*).

g is the output function that calculates the validity degrees of the output CP_{EVAL} as the difference (error) between the calculated WPA and a reference value provided by the therapist (WPA_{ref}).

$$\text{error} = |WPA_i - WPA_{ref}| \tag{6}$$

This error is fuzzified using linguistic labels defined by their vertices as follows: $\{P\ (0, 0, 0.11C_1, 0.25C_1)$, $Z\ (0.11C_1, 0.25C_1, 0.46C_1, 0.6C_1)$, $N\ (0.46C_1, 0.6C_1, \infty, \infty)\}$. Note that we have reused the scale factor C_1 to represent this error measure.

T produces linguistic expressions as follows: *"The evaluation is* {*negative* | *zero* | *positive*}*"*.

4.2. Activation of the avatar (PM_{ACTIV})

This PM allows us to determine the emotional intensity of the messages or advices. It has the following elements:

U is a vector with the numerical value of WPA and the user emotional stability.

y is the output CP_{ACTIV}, where A_y =(*Passive (P), Zero (Z), Active (A)*).

g is the output function that calculates the validity degrees of the output CP_{ACTIV}. We calculate the difference between the current error ($error[t]$) and the error calculated the previous week ($error[t-1]$), as follows:

$$\Delta error = error[t] - error[t-1] \quad (7)$$

When $\Delta error$ has positive sign (worsening), the avatar has an *active*, more energetic, activation; when both errors are similar (maintaining), the activation is *zero*; and, finally, when $\Delta error$ has negative sign (improving), the avatar adopts a *passive*, more relaxed, activation.

In addition, we include as input the type of emotional stability of the user. This allows to change the model depending on whether the user is emotionally stable or unstable. According to the user's pathology, this mechanism allows the designer to tune the way in which the avatar communicates the results. The linguistic labels of the activation are defined as follows: $\{P\ (-\infty, -\infty, -0.2C_2, -0.1C_2)$, $Z\ (-0.2C_2, -0.1C_2, 0.1C_2, 0.2C_2)$, $A\ (0.1C_2, 0.2C_2, \infty, \infty)\}$, where C_2 is a fuzzy factor established by the therapists to indicate the emotional stability degree of the user ($0 < C_2 \leq 1$). The higher C_2, the more unstable is the user. Thus, for example, if a user is emotionally unstable, we increase the support of the linguistic label *Zero*, smoothing the activation level and avoiding extreme emotional states such as *Angry* or *Excited*.

T produces linguistic expressions as follows: *"The activation is {passive | zero | active}"*.

4.3. Emotional state of the virtual agent ($PM_{EMOTION}$)

This PM provides the emotional state of the virtual agent. It has the following elements:

U is the vector: $\{CP_{EVAL}, CP_{ACTIV}\}$.

y is the output $CP_{EMOTION}$, where A_y =(*Neutral* (q_1), *Worried* (q_2), *Excited* (q_3), *Happy* (q_4), *Satisfied* (q_5), *Hopeful* (q_6), *Disgusted* (q_7), *Sad* (q_8), *Angry* (q_9)).

g is the aggregation function that calculates the validity degrees of the output $CP_{EMOTION}$. Here, we use a FFSM that calculates the emotional state based on the previous emotional state and the current inputs CP_{EVAL} and CP_{ACTIV}. This is a characteristic that makes that our emotional model works better than traditional ones since the inertia given by the previous state avoid jumps between extreme states, which is an undesirable condition when you want to simulate the mood of a virtual agent as real as possible. For example, to directly change from *angry* state to *satisfied* state is unrealistic even though evaluation and activation conditions are fulfilled, and may cause the user to feel confused.

In order to implement a specific motivation strategy, we have carefully chosen the transitions among the emotional states according to the diagram shown in Fig. 6. The FFSM is composed by the following set of rules:

R_{11}: IF (*S[t]* is *neutral*) AND (*eval.* is *zero*) AND (*activ.* is *zero*) THEN (*S[t+1]* is *neutral*)

R_{12}: IF (*S[t]* is *neutral*) AND (*eval.* is *zero*) AND (*activ.* is *active*) THEN (*S[t+1]* is *worried*)

R_{13}: IF (*S[t]* is *neutral*) AND (*eval.* is *positive*) THEN (*S[t+1]* is *excited*)

R_{16}: IF (*S[t]* is *neutral*) AND (*eval.* is *zero*) AND (*activ.* is *passive*) THEN (*S[t+1]* is *hopeful*)

R_{17}: IF (*S[t]* is *neutral*) AND (*eval.* is *negative*) THEN (*S[t+1]* is *disgusted*)

R_{21}: IF (*S[t]* is *worried*) AND (*eval.* is *zero*) AND (*activ.* is *passive*) THEN (*S[t+1]* is *neutral*)

R_{22}: IF (*S[t]* is *worried*) AND (*eval.* is *zero*) AND (*activ.* is *zero* OR *activ.* is *active*) THEN (*S[t+1]* is *worried*)

R_{23}: IF (*S[t]* is *worried*) AND (*eval.* is *positive*) THEN (*S[t+1]* is *excited*)

R_{27}: IF (*S[t]* is *worried*) AND (*eval.* is *negative*) THEN (*S[t+1]* is *disgusted*)

R_{31}: IF (*S[t]* is *excited*) AND (*eval.* is *negative*) THEN (*S[t+1]* is *neutral*)

R_{33}: IF (*S[t]* is *excited*) AND (*eval.* is *positive*) AND (*activ.* is *zero* OR *activ.* is *active*) THEN (*S[t+1]* is *excited*)

R_{34}: IF (*S[t]* is *excited*) AND (*eval.* is *positive*) AND (*activ.* is *passive*) THEN (*S[t+1]* is *happy*)

R_{36}: IF (*S[t]* is *excited*) AND (*eval.* is *zero*) THEN (*S[t+1]* is *hopeful*)

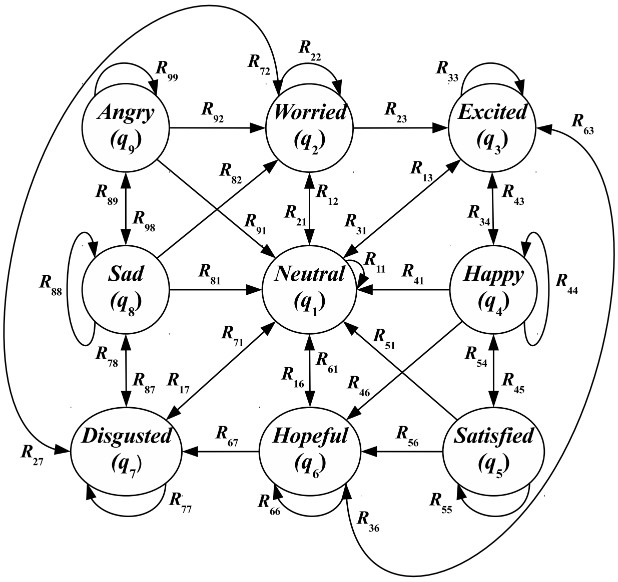


Figure 6: States diagram of the FFSM.

R_{41}: IF (*S[t]* is *happy*) AND (*eval.* is *negative*) THEN (*S[t+1]* is *neutral*)
R_{43}: IF (*S[t]* is *happy*) AND (*eval.* is *positive*) AND (*activ.* is *active*) THEN (*S[t+1]* is *excited*)
R_{44}: IF (*S[t]* is *happy*) AND (*eval.* is *positive*) AND (*activ.* is *zero*) THEN (*S[t+1]* is *happy*)
R_{45}: IF (*S[t]* is *happy*) AND (*eval.* is *positive*) AND (*activ.* is *passive*) THEN (*S[t+1]* is *satisfied*)
R_{46}: IF (*S[t]* is *happy*) AND (*eval.* is *zero*) THEN (*S[t+1]* is *hopeful*)
R_{51}: IF (*S[t]* is *satisfied*) AND (*eval.* is *negative*) THEN (*S[t+1]* is *neutral*)
R_{54}: IF (*S[t]* is *satisfied*) AND (*eval.* is *positive*) AND (*activ.* is *active*) THEN (*S[t+1]* is *happy*)
R_{55}: IF (*S[t]* is *satisfied*) AND (*eval.* is *positive*) AND (*activ.* is *passive* OR *activ.* is *zero*) THEN (*S[t+1]* is *satisfied*)
R_{56}: IF (*S[t]* is *satisfied*) AND (*eval.* is *zero*) THEN (*S[t+1]* is *hopeful*)
R_{61}: IF (*S[t]* is *hopeful*) AND (*eval.* is *zero*) AND (*activ.* is *active*) THEN (*S[t+1]* is *neutral*)
R_{63}: IF (*S[t]* is *hopeful*) AND (*eval.* is *positive*) THEN (*S[t+1]* is *excited*)
R_{66}: IF (*S[t]* is *hopeful*) AND (*eval.* is *zero*) AND (*activ.* is *passive* OR *activ.* is *zero*) THEN (*S[t+1]* is *hopeful*)
R_{67}: IF (*S[t]* is *hopeful*) AND (*eval.* is *negative*) THEN (*S[t+1]* is *disgusted*)
R_{71}: IF (*S[t]* is *disgusted*) AND (*eval.* is *positive*) THEN (*S[t+1]* is *neutral*)
R_{72}: IF (*S[t]* is *disgusted*) AND (*eval.* is *zero*) THEN (*S[t+1]* is *worried*)
R_{77}: IF (*S[t]* is *disgusted*) AND (*eval.* is *negative*) AND (*activ.* is *passive* OR *activ.* is *zero*) THEN (*S[t+1]* is *disgusted*)
R_{78}: IF (*S[t]* is *disgusted*) AND (*eval.* is *negative*) AND (*activ.* is *active*) THEN (*S[t+1]* is *sad*)
R_{81}: IF (*S[t]* is *sad*) AND (*eval.* is *positive*) THEN (*S[t+1]* is *neutral*)
R_{82}: IF (*S[t]* is *sad*) AND (*eval.* is *zero*) THEN (*S[t+1]* is *worried*)
R_{87}: IF (*S[t]* is *sad*) AND (*eval.* is *negative*) AND (*activ.* is *passive*) THEN (*S[t+1]* is *disgusted*)
R_{88}: IF (*S[t]* is *sad*) AND (*eval.* is *negative*) AND (*activ.* is *zero*) THEN (*S[t+1]* is *sad*)
R_{89}: IF (*S[t]* is *sad*) AND (*eval.* is *negative*) AND (*activ.* is *active*) THEN (*S[t+1]* is *angry*)
R_{91}: IF (*S[t]* is *angry*) AND (*eval.* is *positive*) THEN (*S[t+1]* is *neutral*)
R_{92}: IF (*S[t]* is *angry*) AND (*eval.* is *zero*) THEN (*S[t+1]* is *worried*)
R_{98}: IF (*S[t]* is *angry*) AND (*eval.* is *negative*) AND (*activ.* is *passive*) THEN (*S[t+1]* is *sad*)
R_{99}: IF (*S[t]* is *angry*) AND (*eval.* is *negative*) AND (*activ.* is *zero* OR *activ.* is *active*) THEN (*S[t+1]* is *angry*)

The firing degree of each rule is calculated like in PM_{STATE}.

T produces linguistic expressions as follows: *"The emotional agent is* $\{$*neutral* | *hopeful* | *excited* | *happy* | *satisfied* | *worried* | *disgusted* | *sad* | *angry*$\}$*"*.

5. Report Template

The Expression Module, according to the Report Template, allows the designer to analyze the $GLMP_{Activity}$ and highlight, according to the $GLMP_{Emotion}$, the most interesting features for users. In the application described in this paper, the Expression Module is able to provide the linguistic reports explained in the following subsections:

5.1. Daily reports

These linguistic reports deal with allowing the user to self-track his/her daily physical activity. The daily report starts with the information obtained from the CP_{DPA}:

"Today, you had a {sedentary | moderate | active | vigorous | extreme} lifestyle."

In addition, according to the PM_{CR}, the model provides the amount of calories that the user needs intake to get a healthy physical state. This information could be relevant in those applications that deal with physical disorders, e.g., obesity or anorexia, but in those that deal with other type of disorders could be less interesting. Thus, the therapist decides if the following information will be presented to the user in the final report:

"According to this lifestyle and your BMI, you must intake {much less than | less than | approximately | more than | much more than} "x" kcal. to {loss | maintain | gain} weight."

5.2. Weekly reports

These linguistic reports describe the user lifestyle during a week and they include emotional advices that vary according to the obtained results. As showed in Fig. 7, the template of the weekly reports is structured with the drawing of the avatar's facial expression according to its emotional state and a balloon text with three different parts, namely, *Greeting*, *Lifestyle* and *Emotional advice*.

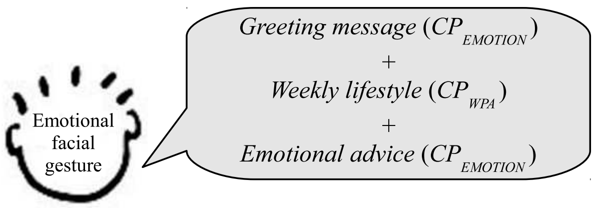


Figure 7: Report template for the weekly reports.

Greeting: the content of the emotional greeting is based on the information extracted from the $CP_{EMOTION}$. Thus, for example, if the emotional state is *excited*, the greeting could be "*Great!*"; but if the avatar is *disgusted*, the greeting would be different than the previous one, e.g., "*Hey, what happened? Your physical activity is still getting worse. You have not heeded my advices.*"

Lifestyle: the middle part of the message informs about the current lifestyle based on the results reported by the CP_{WPA}.

Emotional advice: finally, the message includes an emotional advice that congratulates the user when the results are good and motivates him/her to improve the physical activity and reach the goals. Examples of these emotional advices are "*You are doing great!*" or "*I am sure that you could do it better, come on!*".

6. Experimentation

This section presents the two experiments we have performed to demonstrate the potential and applicability of our approach. This experimentation could be considered only the first stage for testing the viability and possibilities of a future complete and functional tool.

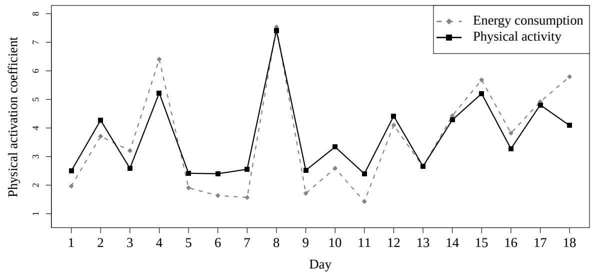


Figure 8: Energy consumption vs. Physical activity.

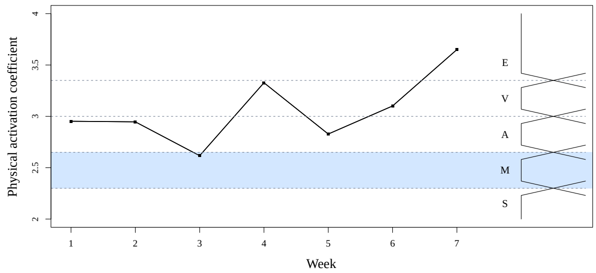


Figure 9: WPA evolution (P.1).

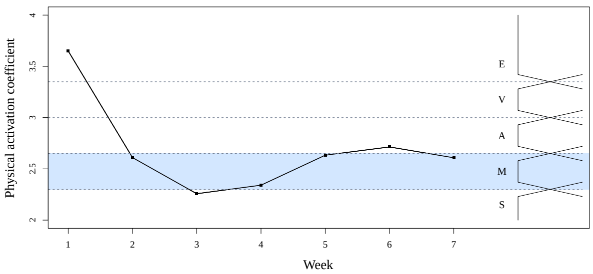


Figure 10: WPA evolution (P.2).

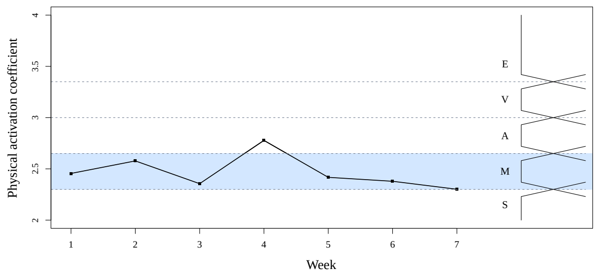


Figure 11: WPA evolution (P.3).

6.1. Estimation of the physical activity

Objective. To verify that our way of estimating the physical activity is correct, contrasting the calculated physical activity with the energy consumption measured by using a commercial heart rate monitor.

Experimental layout. The test procedure consisted of monitoring the physical activity of an 80 kilograms man while carrying two measurement devices simultaneously. The first one was a smartphone that the user placed in different locations (breast pocket, trousers pocket, armband, backpack, ...) during a period of 18 days. The acquisition of the accelerations was done with the Android application explained in Section 2. The second one was a commercial heart rate monitor fixed in his breast that estimates the energy consumption by combining the heart rate with the users' gender, weight and age.

Results. Fig. 8 shows a graphical representation of the evolution of the physical activity versus the energy consumption (in kcal/min) monitored by the heart rate monitor during 18 days.

Discussion of results. The measurement provided by the heart rate monitor is very reliable since it takes into account the physical condition of the user, reflected in his heart rate.

Results reveal a close correlation between the physical activity estimated with our model and the energy consumption offered by the heart rate monitor (the Pearson correlation coefficient is equal to 0.93). As we can appreciate in Fig. 8, the trend is exactly the same except for the days 7 and 18, when the energy consumption had a different direction from the estimated physical activity. Therefore, we considered that monitoring the physical activity of users is enough to understand changes and maintain exercise routines without the need of additional devices.

6.2. Self-tracking the physical activity

Objective. In order to verify the evolution of the emotional state of the designed avatar, we performed an experiment based on tracking the physical activity of three subjects for a period of seven weeks. This experiment reproduces the physical behavior of a person that visits a therapist and is required to self-track his/her physical activity to reach some goals, e.g. losing weight in a controlled manner or improving the

	Week	CP_{EVAL}			CP_{ACTIV}			$CP_{EMOTION}$								
		N	**Z**	**P**	**P**	**Z**	**A**	**N**	**W**	**E**	**Ha**	**S**	**H**	**D**	**Sa**	**A**
P.1	1	0.12	0.88	0	0	1	0	0	0	0	0	0	**0.88**	0.12	0	0
	2	0.08	0.92	0	0	1	0	0	0.11	0	0	0	**0.82**	0.07	0	0
	3	0	0.23	0.77	1	0	0	0.06	0.02	**0.71**	0	0	0.21	0	0	0
	4	1	0	0	0	0	1	**0.77**	0	0	0	0	0	0.23	0	0
	5	0	1	0	1	0	0	0	0.23	0	0	0	**0.77**	0	0	0
	6	1	0	0	0	0	1	0	0	0	0	0	0	**1**	0	0
	7	1	0	0	0	0	1	0	0	0	0	0	0	0	**1**	0
P.2	1	1	0	0	0	1	0	0	0	0	0	0	0	**1**	0	0
	2	0	0.18	0.82	1	0	0	**0.82**	0.18	0	0	0	0	0	0	0
	3	0	0.77	0.23	0	0.36	0.64	0.27	**0.49**	0.23	0	0	0	0	0	0
	4	0	0.17	0.83	0.67	0.33	0	0.09	0.04	**0.62**	0.19	0	0.06	0	0	0
	5	0	0.35	0.65	0	1	0	0.04	0.02	**0.51**	0.16	0	0.27	0	0	0
	6	0	0.92	0.08	0	0.39	0.61	0.22	0.03	0.06	0.01	0	**0.68**	0	0	0
	7	0	0.17	0.83	1	0	0	0.01	0	**0.77**	0.07	0.01	0.15	0	0	0
P.3	1	0	0	1	0	1	0	0	0	0	**1**	0	0	0	0	0
	2	0	0	1	0	0.34	0.66	0	0	**0.66**	0.34	0	0	0	0	0
	3	0	0.07	0.93	0	1	0	0	0	**0.63**	0.32	0	0.05	0	0	0
	4	0	1	0	0	0	1	0.07	0	0	0	0	**0.93**	0	0	0
	5	0	0	1	1	0	0	0	0	**1**	0	0	0	0	0	0
	6	0	0	1	0	1	0	0	0	**1**	0	0	0	0	0	0
	7	0	0.46	0.54	0	0.44	0.56	0	0	**0.54**	0	0	0.46	0	0	0

Table 1: Emotional states of the avatar during the 5 monitored months.

quality of life. To help the user to reach these objectives, our computational tool provides a daily linguistic report that summarizes his/her physical activity during the day and makes a recommendation about the energy requirement needed to get a healthy physical state. The avatar motivates, advises or scolds the user, communicating the obtained results with emotional emphasis.

Experimental layout. The experiment consisted on analyzing the accelerations collected by the smartphone of each participant (P.1, P.2 and P.3) during a period of seven weeks. Users started the application in the morning and stopped it at night, when data were sent to a central server and were analyzed. With the received whole day data we applied the $GLMP_{Activity}$ and estimated the users' daily physical activity as explained in Section 3.4. When the week finished, we calculated the users' weekly physical activity as explained in Section 3.6.

Results. Figs. 9, 10 and 11 show the evolution of the weekly physical activity over the seven monitored weeks. We can see in dashed lines the cutoffs of each linguistic label associated to the weekly physical activity (CP_{WPA}).

We have simulated that, for this experiment, users had the goal to reach and maintain a "moderate" lifestyle. Following these premises, Table 1 specifies, for each of the analyzed weeks, which were the evaluation and activation results and, in consequence, which was the emotional state of the avatar after applying the computational tool explained in Section 4. Each emotional state is represented as follows: Neutral (N), Worried (W), Excited (E), Happy (Ha), Satisfied (S), Hopeful (H), Disgusted (D), Sad (Sa) and Angry (A).

Discussion of results. We can see how, depending on the users' weekly physical activity shown in Figs. 9, 10 and 11, the avatar adopts an emotional state to helps the user to reach the goal established by the therapist, i.e., to get a "moderate" lifestyle. The three participants of this experiment had a different evolution during the analyzed seven weeks, which is explained as follows:

- Participant 1 started the experimental phase with an "active" lifestyle which made the avatar be *hopeful* of achieving the objective soon. Following two weeks confirmed this good trend, making that the avatar felt *excited*. However, this situation was reversed and, in the following weeks, the user progressively increased her physical activity, disregarding the avatar advices. This situation causes that the avatar looked *disgusted* and, finally, *sad* with the user's behavior.

- Participant 2, unlike participant 1, started the experimental phase with an "extreme" lifestyle, too far from being a "moderate" lifestyle, causing that the avatar looked *disgusted* at this stage. The user change the physical behavior and next week, he reduced the physical activity to a "moderate" level. The avatar, despite this abrupt change, looked with a *neutral* emotion. This is one of the advantages of our emotional model, i.e., including previous emotion in the emotional model avoids jumps between extreme states, although conditions indicate that the objective has been achieved. This situation makes the avatar to look more real, like a human would look, being suspicious of whether this change would be coincidental or would not. As in the following weeks this trend was, more or less, maintained, the avatar finally looked *excited*.

- Participant 3, meanwhile, started the experimental phase with a "moderate" lifestyle which made the avatar to feel *happy*. Next weeks, the user was balancing around this level, so the predominant avatar's emotional state was *excited*. In the fourth week, the user had an "active" lifestyle and the avatar looked hopeful that the user return to the objective level.

7. Conclusions

We have organized our research around the practical project of developing a new type of computational application able to monitor the activities of users and communicate them by using natural language messages that include emotional content. This challenge has led us to tackle several contributions summarized as follows:

- We have obtained a feasible measure of the physical activity without the requirement of carrying the smartphone in specific locations in the body.

- Versus normally black box approaches, we have designed an interpretable and transparent model of the basic human physical activity that uses natural language to describe patterns emerging in data.

- We have designed a virtual agent with a simulated emotional state that evolves depending on input data and generates meaningful reports and advices.

Concepts like self-tracking and quantified self are starting to be generally known. Currently, there is a clear demand of new tools that will be able to help therapists and users to address the treatment of pathologies in a more rapid and effective way than the existing ones.

This is only the first step of an ambitious project that will require multidisciplinary collaboration with physiotherapists and psychologists. The results presented in this paper demonstrate the viability of this type of computational systems. As stated in the Introduction, future research will conduct to demonstrate the effectiveness of these emotional messages to motivate changes in the physical behavior of people. In addition, we are evaluating the possibility of developing a commercial application that uses the information provided by wearable technologies, such as smart wristbands, which are becoming a powerful entry in the market of smart technologies.

Acknowledgment

This work has been funded by the Spanish Government (MICINN) under the project TIN2011-29827-C02-01 and the Principality of Asturias Government under the project CT13-52.

References

Alepis, E., Virvou, M., 2011. Automatic generation of emotions in tutoring agents for affective e-learning in medical education. Expert Systems with Applications 38, 9840–9847.

Alsina, C., Frank, M.J., Schweizer, B., 2006. Associative Functions: Triangular Norms and Copulas. World Scientific Publishing Company, Singapore.

A. Alvarez-Alvarez, D. Sanchez-Valdes, G. Trivino, A. Sánchez, P. D. Suárez, Automatic linguistic report about the traffic evolution in roads, Expert Systems with Applications 39 (2012) 11293–11302.

Alvarez-Alvarez, A., Alonso, J.M., Trivino, G., 2013. Human activity recognition in indoor environments by means of fusing information extracted from intensity of wifi signal and accelerations. Information Sciences 233, 162–182.

Alvarez-Alvarez, A., Trivino, G., 2013. Linguistic description of the human gait quality. Engineering Applications of Artificial Intelligence 26, 13–23.

Alvarez-Alvarez, A., Trivino, G., Cordón, O., 2012. Human gait modeling using a genetic fuzzy finite state machine. Fuzzy Systems, IEEE Transactions on 20. doi:10.1109/TFUZZ.2011.2171973.

Bailador, G., Trivino, G., 2010. Pattern recognition using temporal fuzzy automata. Fuzzy Sets and Systems 161, 37–55.

Bao, L., Intille, S.S., 2004. Activity recognition from user-annotated acceleration data, in: Proceedings of the 2nd International Conference on Pervasive Computing, pp. 1–17.

Booth, M.L., Ainsworth, B.E., Pratt, M., Ekelund, U., Yngve, A., Sallis, J.F., Oja, P., 2003. International physical activity questionnaire: 12-country reliability and validity. Medicine & Science in Sports & Exercise 195, 3508–1381.

Damasio, A., 2008. Descartes' error: Emotion, reason and the human brain. Random House.

El-Nasr, M.S., Yen, J., Ioerger, T.R., 2000. Flamefuzzy logic adaptive model of emotions. Autonomous Agents and Multi-agent systems 3, 219–257.

FAO/WHO-OMS/UNU, 1985. Expert consultation report. Energy and Protein Requirements. Technical Report Series 724.

Fogg, B.J., 2002. Persuasive technology: using computers to change what we think and do. Ubiquity 2002, 5.

Fragopanagos, N., Taylor, J.G., 2005. Emotion recognition in human–computer interaction. Neural Networks 18, 389–405.

van der Heide, A., Trivino, G., 2010. Simulating emotional personality in human computer interfaces, in: Fuzzy Systems (FUZZ), 2010 IEEE International Conference on, IEEE. pp. 1–7.

Kenny, P.G., Parsons, T.D., Rizzo, A.A., 2009. Human computer interaction in virtual standardized patient systems, in: Human-Computer Interaction. Interacting in Various Application Domains. Springer, pp. 514–523.

Kwapisz, J., Weiss, G., Moore, S., 2011. Activity recognition using cell phone accelerometers. ACM SIGKDD Explorations Newsletter 12, 74–82.

Kwon, Y., Kang, K., Bae, C., 2014. Unsupervised learning for human activity recognition using smartphone sensors. Expert Systems with Applications 41, 6067–6074.

Mahamood, S., Reiter, E., 2011. Generating affective natural language for parents of neonatal infants, in: Proceedings of the 13th European Workshop on Natural Language Generation, Association for Computational Linguistics. pp. 12–21.

Martınez-Miranda, J., Aldea, A., 2005. Emotions in human and artificial intelligence. Computers in Human Behavior 21, 323–341.

Mifflin, M.D., St Jeor, S., Hill, L.A., Scott, B.J., Daugherty, S.A., Koh, Y., 1990. A new predictive equation for resting energy expenditure in healthy individuals. The American journal of clinical nutrition 51, 241–247.

Ouchi, K., Doi, M., 2012. Living activity recognition using off-the-shelf sensors on mobile phones. Annals of Telecommunications , 1–9.

Picard, R.W., 2000. Affective computing. MIT press.

Plutchik, R., 1980. Emotion: A psychoevolutionary synthesis. Harper & Row New York.

Reiter, E., Dale, R., 2000. Building natural language generation systems. volume 152. MIT Press.

Robertson, W., Stewart-Brown, S., Wilcock, E., Oldfield, M., Thorogood, M., 2011. Utility of accelerometers to measure physical activity in children attending an obesity treatment intervention. Journal of Obesity .

Ruspini, E.H., 1969. A new approach to clustering. Information and Control 15, 22–32.

Russell, J.A., 1979. Affective space is bipolar. Journal of personality and social psychology 37, 345.

Sanchez-Valdes, D., Alvarez-Alvarez, A., Trivino, G., Linguistic description about circular structures of the Mars' surface, Applied Soft Computing 13 (2013) 4738–4749.

Sanchez-Valdes, D., Trivino, G., 2013. Computational perceptions of uninterpretable data. A case study on the linguistic modeling of human gait as a quasi-periodic phenomenon. Fuzzy Sets and Systems 253 (2014) 101–121.

Unold, O., et al., 2011. Machine learning approach to model sport training. Computers in human behavior 27, 1499–1506.

Whissell, C., 1989. The dictionary of affect in language. Emotion: Theory, research, and experience 4, 94.

WHO, W.H.O., 2004. Appropriate body-mass index for asian populations and its implications for policy and intervention strategies. The Lancet 10, 157–163.

Zadeh, L.A., 1975. The concept of a linguistic variable and its application to approximate reasoning. Parts I, II, and III. Information Sciences 8, 8, 9, 199–249, 301–357, 43–80.

Zadeh, L.A., 1999. From computing with numbers to computing with words - from manipulation of measurements to manipulation of perceptions. Circuits and Systems - I: Fundamental theory and applications, IEEE Transactions on 45(1), 105–119.

Zadeh, L.A., 2008. Toward human level machine intelligence - is it achievable? the need for a paradigm shift. IEEE Computational Intelligence Magazine .