

Introducing Computational Semantics for Natural Language Understanding in Conversational Nutrition Coaches for Healthy Eating [†]

Antonio Benítez-Guijarro ^{*}, Zoraida Callejas, Manuel Noguera and Kawtar Benghazi

Department of Languages and Computer Systems, University of Granada, 18071 Granada, Spain; zoraida@ugr.es (Z.C.); mnoguera@ugr.es (M.N.); benghazi@ugr.es (K.B.)

^{*} Correspondence: ajbenitez@ugr.es; Tel.: +34-958-241000 (ext. 20049)

[†] Presented at the 12th International Conference on Ubiquitous Computing and Ambient Intelligence (UCAmI 2018), Punta Cana, Dominican Republic, 4–7 December 2018.

Published: 18 October 2018

Abstract: Nutrition e-coaches have demonstrated to be a successful tool to foster healthy eating habits, most of these systems are based on graphical user interfaces where users select the meals they have ingested from predefined lists and receive feedback on their diet. On one side the use of conversational interfaces based on natural language processing allows users to interact with the coach more easily and with fewer restrictions. However, on the other side natural language introduces more ambiguity, as instead of selecting the input from a predefined finite list of meals, the user can describe the ingests in many different ways that must be translated by the system into a tractable semantic representation from which to derive the nutritional aspects of interest. In this paper, we present a method that improves state-of-the-art approaches by means of the inclusion of nutritional semantic aspects at different stages during the natural language understanding processing of the user written or spoken input. The outcome generated is a rich nutritional interpretation of each user ingest that is independent of the modality used to interact with the coach.

Keywords: natural language processing; nutrition; E-health; voice assistant; semantic processing; multimodal interaction

1. Introduction

Nowadays, there is an upsurge of the development and use of applications aimed at controlling and fostering healthy eating habits [1]. These virtual nutritional coaches usually manage the amount of recommended nutrients through a meal record system [2]. Such record mechanism is based on the direct interaction of users through predefined lists of dish recipes for the selection of the ingested meals. However, although this is a useful method, it is also tedious, as users experience difficulties interacting with these tools since the process of recording the meals is usually very time consuming. Additionally, the use of predefined lists for meal registration requires that, for complex compound meals that are not preregistered in the list, nutritional information must be calculated manually. For instance, the calories of a portion of margherita pizza would have to be computed considering the calories of the proportion of the ingredients that compose it, i.e., the calories corresponding to the amount of cheese, flour, tomato and rest of ingredients. An alternative is to develop applications with natural language interfaces with which users can interact orally or through written inputs [3,4]. The current solutions for nutritional coaching that make use of natural language processing are mainly based on the detection of keywords [5]. After analyzing the keywords, these applications identify the action indicated and execute a response in the form of a conversation or reaction. The

main drawback in these methods is that the inputs that match the desired action are quite limited and they lack of flexibility. Thus, conversations between users and applications are far from being natural, which increases reluctance and confidence in the virtual coach [6].

This paper presents a proposal that seeks to improve the current state of the art related to interaction between nutritional coaching software systems and their users. The proposal presented goes a step further in the application of natural language processing in assistants and nutritional coaching by introducing a mechanism based on the syntactic and semantic analysis of sentences instead of using conventional keyword spotting. The proposal is based on the analysis of syntagmas that contain valuable information for computing meal records through the syntactic parsing of user's commands [7]. In our approach, the phrases are parsed into syntagmas that can be classified into categories from the nutritional domain. The syntactic parsing also reveals the syntactic role and dependencies of syntagmas, which we consider to interpret their meaning. The result of this process is a semantic tree that includes the dependencies detected. The tree is interpreted against a general structure of logical predicates [7–9] that we have defined for the field of nutritional recording. This general structure has been proposed from a detailed study about the most relevant natural language commands used to record meals.

With this technique, it is possible to obtain more meaningful and complex interpretations than those currently achieved with keyword spotting. As we consider the structure of the phrases, we can analyse the existing dependencies. For example, "an omelet with two eggs" can only be interpreted correctly when it is clear that "two eggs" complements the omelet. Systems based on keyword spotting would have interpreted that the user has eaten eggs and omelet (not an omelet cooked with two eggs), and there would be an ambiguity about how to assign the quantity (two eggs of two omelets). Our evaluation results confirm that the proposed approach obtains higher interpretation accuracy and makes it possible to interpret more complex commands (e.g., compound meals).

The rest of the paper is organized as follows. Section 2, presents the current approaches for language understanding in the nutritional coaches field. Section 3 defines a new natural language interaction procedure for nutritional coaches to improve the food recording process. Section 4 shows the general discussion behind the validation process and the results. Finally, Section 5 presents the conclusions.

2. Current Approaches for Language Understanding in Nutritional Coaches

In the following section we introduce the general background in which this proposal is located. To describe the general context, we are going to talk about the health coaching and the current advances in the use of semantic rules in the field of nutritional coaching.

2.1. Nutritional Coaching

Health coaching is based on the sustained review of different healthy habits such as nutrition, physical exercise, sleep habits, etc. [10,11]. This discipline aims to identify and correct healthy behaviors with a pre-established goal. Among the different areas that it encompasses, we highlight nutritional coaching. Nutritional coaching is the training and supervision of the nutritional habits of users with the objective of adapting their alimentary intakes to meet health goals [12]. To perform this supervision, an analysis process of food consumed by users is necessary. There exist several mechanisms to track this supervision like following a strict diet plan, annotating the usual eaten meals or completing nutritional surveys in order to know the user's nutritional habits [13]. Nowadays, the incursion of new technologies has favored the use of software to monitor nutritional analysis [14]. These tools are used by users to search and record daily food intakes in order to subsequently perform an analysis of the nutritional information of the food eaten [15]. With this analysis the dietary habits of the monitored users are cataloged and nutritional plans (diets) can be suggested to each user.

2.2. Semantics for Health Coaches

Language is the main mechanism for communication and analysis of information. Therefore, through natural language processing (NLP), we seek to find computational mechanisms that allow us to understand and generate texts in a natural language automatically. With the introduction of new interaction mechanisms through natural language, we can provide software systems with more natural methods to interact with users. To perform automatic natural language understanding, it is necessary to extract and identify all the structures with value that are going to be used (e.g., “I have eaten three red and tasty apples” the value structures for the processing are “have eaten” that indicate the action, “three apples” that indicate the main argument which is going to be recorded). For this, different techniques must be used to analyze and process the language of the users. In the context of health coaching, the analysis of key words in natural language is one of the most widely used techniques. Different assistants from the field of health and from other areas use this type of language analysis due with this method it is possible to develop easily new applications that respond to certain specific commands through voice or text commands. Exemplifying this process, there are tools such as DialogFlow that allow the development of assistants or coaches with voice interaction or text entry [16]. For its operation, different fixed or slightly dynamic structures organized by keywords are specified. These keywords are placed in specific positions of the different commands that want to be implemented. During the analysis of natural language, these tools divide the text into pieces analyzing whether the words fit into the previously defined categories. Finally, after the analysis of the evaluated pieces, the action and the arguments to develop it are recognized and the action is executed with that information. These procedures offer effective results but are limited to a preset structure of previously evaluated commands.

Another alternative to process natural language is the syntactic and semantic analysis of the sentences [8]. Syntactic analysis starts with the breakdown of the sentences into their syntactic constituents, identifying their dependencies and relationships. After syntactic analysis, the meaning of the identified syntagmas is analyzed. There are different ways to perform semantic analysis, for example by using statistical methods [17]. We have decided to apply semantic rules to have a means to check the structure of the parsed phrases and detect missing pieces of data or ill-formed structures that are not interpretable in a meal record system. To do that, we have defined a rule system based on logical predicates. As discussed before, these combined processes make it possible to interpret more complex phrases with a greater accuracy than using keyword spotting.

3. Our Approach to Nutritional Computational Semantics

In this section we present the basic concepts on which the research is based, as well as the operation of the mechanism proposed for semantic analysis in the nutritional coaching domain.

In this paper we propose a new natural language interaction method for nutritional coaches that allows describing the meals ingested with dynamic description in order to improve the meals record system used in currently nutrition coaches software. The process of recording meals through instructions in natural language requires being able to identify all possible structures when describing what foods have been eaten. Users describe food eaten in different ways by interacting orally or writing sentences such as: “I have had two toasts and a glass of milk for breakfast”, “I have eaten two pieces of pepperoni pizza”, “For lunch I had rice with chicken and two fried eggs”.

To perform an analysis of the extracted text of the users, it is necessary to identify a general structure where to detect the complements and the main arguments of the sentences. To carry out this process we have analyzed and decomposed different sentences used by users extracting a common structure and some basic “entities” that fulfill all the sentences in this area [17–19]. In order to recognize the descriptions of meals in natural language, we have extracted a series of rules described in first order logic predicates obtained after the transformation of the different syntagmas in the syntactic analysis of the sentences, giving rise to the general rules used in this study. During the process of studying semantic rules, it is necessary to catalog the common value syntagmas in all sentences, giving rise to entities that encompass the common elements in sentences in this area.

These entities are word families with common ontologies. The general entities “Figure 1” collected after the study are:

- Actions: Words that indicate eating or drinking actions that should generate a nutritional record (e.g., I took, I have eaten, I have drunk).
- Lunch schedule: Words or verb tenses that indicate when these foods have been eaten (e.g., breakfast, lunch, dinner).
- Quantifiers: Words that indicate the amount (countable) of food ingested (e.g., one, two, three).
- Containers/units: Words that indicate containers or portions (e.g., glass, bowl).
- Container/unit quantifiers: Words that indicate the sizes of the containers/units of some specific foods (e.g., a large glass of ..., a small piece of ...).
- Simple food or ingredients: Words that indicate basic food (e.g., apple, egg, water).
- Compound meals or recipes: Words that indicate meals made up of other simple or compound foods (e.g., omelet, paella, cake).

<p><i>Actions</i> ⊢ (<i>action</i>, {<i>eat, drink, take, have ...</i>})</p> <p><i>Simple foods or ingredients</i></p> <p>⊢ (<i>food</i>, {<i>apple, milk, egg, peach ...</i>})</p> <p><i>Compound meals or recipes</i></p> <p>⊢ (<i>cFood</i>, {<i>paella, pizza, bread, omelet ...</i>})</p> <p><i>Quantifiers</i> ⊢ (<i>number</i>, {<i>one, two, three, four ...</i>})</p> <p><i>Containers / units</i></p> <p>⊢ (<i>containersUnits</i>, {<i>glass, bowl, piece, slice ...</i>})</p> <p><i>Containers / units quantifiers</i></p> <p>⊢ (<i>containersUnitsQuantifiers</i>, {<i>small, large, big, medium ...</i>})</p> <p><i>Lunch schedule</i></p> <p>⊢ (<i>foodhour</i>, {<i>breakfast, lunck, dinner, snack ...</i>})</p>

Figure 1. Entities for the semantic rules.

Once the syntagmas are decomposed, the relations and dependencies between them are recognized using syntactic analysis. Once the syntactic parsing is done, each piece analyzed in the syntactic parsing is identified by checking the corresponding entities. After this, each syntactic chunk is verified if the predicate (the semantic rule) explained above is fulfilled. All this process is described in the following subsections.

3.1. Procedure

In this subsection we will explain what is the process proposed for the interpretation of alimentary registries expressed in natural language. It is summarized in “Figure 2”.

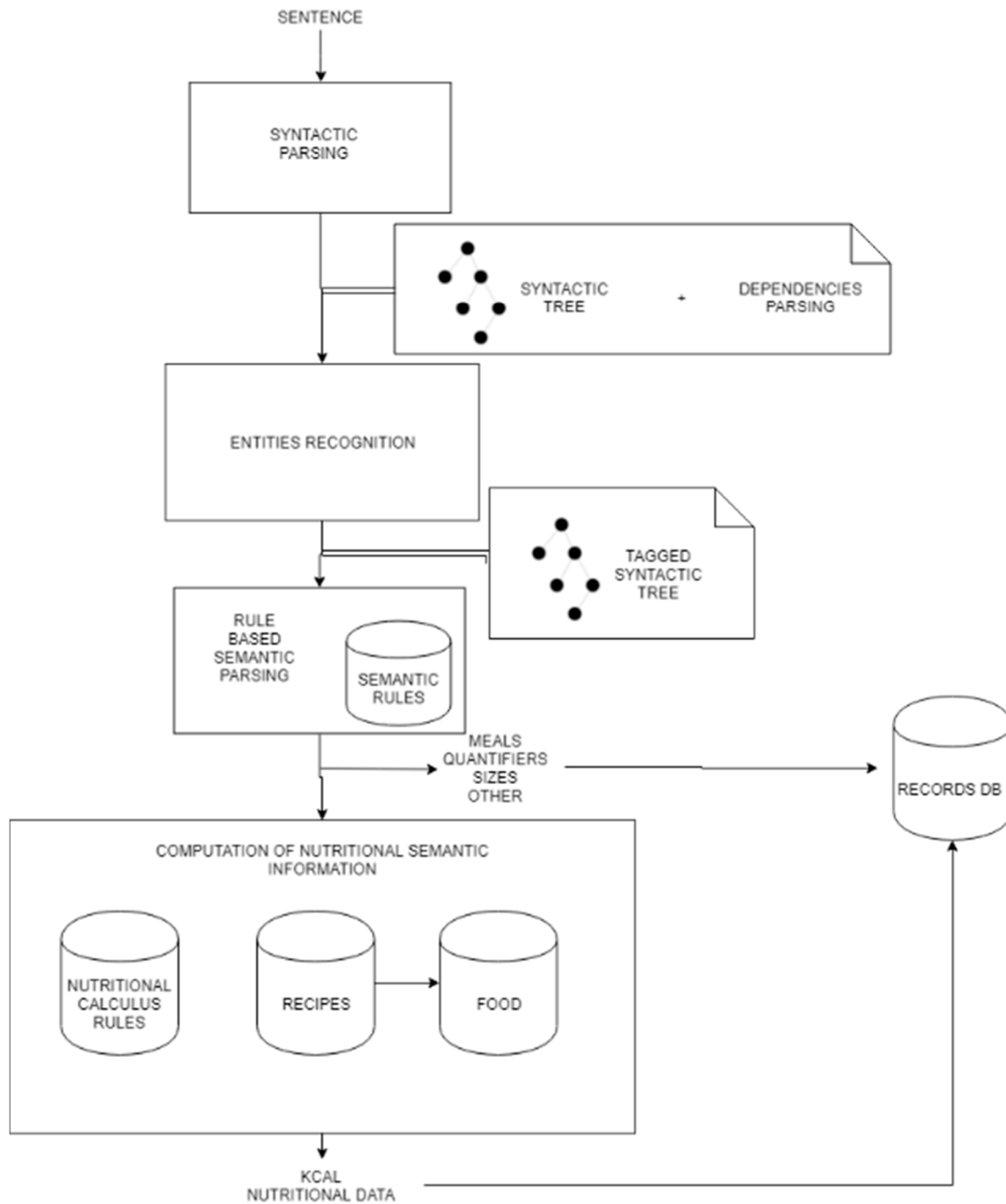


Figure 2. Schema of the proposed procedure.

1. Syntactic parsing

The first stage is to perform syntactic analysis of the user’s input and obtain a syntactic tree with the dependencies and relationships of all elements to be interpreted. In order to do this, state-of-the-art alternatives can be used. There are numerous libraries that provide this functionality such as the Stanford-core-NLP library [20], Apache OpenNLP [21], NLTK [22], etc. The result of the analysis is a tree where the root is the core of the sentence (the verb or action) and the other nodes represent the syntagmas and their relationships.

2. Entities Recognition

After building the syntactic tree and dependencies parsing, the second step is to walk the tree analyzing each node and categorizing it in one of the specific entity groups shown in Figure 1. This process identifies the valuable syntagmas providing all arguments necessary for the nutritional

interpretation. Next, it is necessary to check whether they are compliant with the semantic rules for that node and for the dependencies that hang from it in the form of child nodes (the rules are explained in Section 3.2). The pseudocode for building the tagged syntactic tree is summarized in Figure 3.

Data:
 E ← entities Database (Actions, simple food...)
 St ← Syntactic tree + Dependencies parsing

Result:
 Tt ← Syntactic tree + Dependencies parsing

Semantic tree building Algorithm Semantic tree building based on syntactic parsing

```

1: Root = St.findVerb()
2: If Root ∈ Actions ⊆ then
3:   Build new semantic tree Tt.root = Root
4:   While St.hasChild() do
5:     If St.node() ∈ (E – Actions) then
6:       Tt.node().addChildWithSyntacticDependencies(St.node())
    
```

AddChildWithSyntacticDependencies(node)

```

1: For all syntagmas in node do
2:   If syntagma ∈ (E – Actions) then
3:     nNode = Build new node with syntagma.type
4:     this.addChild(nNode)
    
```

Figure 3. Pseudocode from the entities recognition procedure.

3. Rule based semantic analysis

Once the syntactic tree has been tagged with the most relevant entities, we use a rule-based approach to interpret the meaning of the sentence based on the entities and the dependencies identified by the syntactic parser. This process is described in detail in Section 3.2.

4. Computation of nutritional semantic information

After recognizing each of the necessary arguments and validating them with the rules, they are categorized to calculate the nutritional information. This categorization is performed gathering the information associated with each valuable syntagma from a knowledge base that provides enough information to identify which family of entities each word belongs to (i.e., food information, quantifiers, verbal forms that indicate food intake action and food intake schedules). With a search in the knowledge bases, we complete the ontological information (nutritional information and other fields) of the analyzed value syntagmas.

To associate each argument with nutritional information it is necessary to identify the meals that have been recognized with their quantifiers and then perform all necessary searches in the database to obtain the corresponding nutritional information from the decomposition of the meal into simpler ingredients. This is a major advance of our proposal compared to the current keyword-based approaches. For example, if we analyze the phrase “I have eaten a bowl of milk with cereals and an omelet with three eggs” with both mechanisms the bowl of milk with cereals are recognizable, but the omelet indicates a meal and a supplement (three eggs) that describes its composition. With the processing based on keywords we would not know that the food “omelet” is complemented with the phrase “three eggs” modifying its nutritional composition.

Finally, the last stage is responsible for calculating the nutritional information that will be registered in the system according to all the information gathered from the previous interpretation together with the searches in the database. For this purpose, it will compute an estimation of the

quantity ingested based on quantifiers and size of the containers, which is multiplied by the calories and nutritional information associated with each food. The compound meals are composed of other single food that are measured by a quantifier of portions and its nutritional information. To perform a nutritional calculation of the compound meals the procedure that we propose is to search in the knowledge bases (as it is explained in the previous subsection) to gather information about how to cook the meal. Recipes have the ingredients that compose the meal and their amounts, which are processed as basic food as shown in Figure 4 and aggregated. The proposed procedure enables to alter the composition of the ingredients that compose a compound meal, making it possible for the user to provide more accurate descriptions. For example, if we interpret the sentence “I have eaten an omelet with two eggs and a can of tuna”, the syntactic analysis would be carried out extracting the dependencies and relations of omelet with the foods that describe it (two eggs and a can of tuna). After doing the search in the database, we would extract that “omelet” is a food composed of “eggs”, “oil” and “salt”, and we would interpret that there is an additional ingredient: tuna. After collecting the composition of the omelet from the database and its proportions, the calculation of the nutritional information would be carried out. It would be translated into a mathematic equation in order to perform the nutritional calculus, resulting in: “2” (amount identified in the analysis, if it had not been specified, the pick-up of the omelet recipe would be used, as is the case with the amount of oil and salt) * nutrition information for eggs + nutritional information of “1” can (Xg that makes up the can container) of tuna + Xg (identified in the search of the recipe of omelet in the database) * nutritional information of oil + Xg (identified in the search of the omelet recipe in the database) * salt nutritional information.

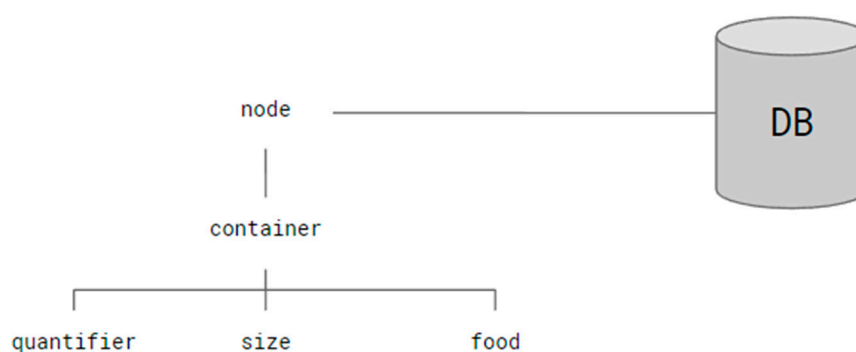


Figure 4. Nutrients calculus step.

3.2. Semantic Rules

As previously mentioned, to implement our proposal it is necessary to carry out a study on the general structure of the natural language phrases that can be employed by the users to to make food records. The general rule for a sentence to be interpretable is represented in (1).

$$\forall x \text{ INTERPRETABLE_PHRASE}(x) \rightarrow \text{EAT}(x) \vee \text{DRINK}(x) \quad (1)$$

The sentences interpretable by the mechanism proposed in this study respond to requests for the registration of food and/or beverages. Therefore, this simplified equation will be fulfilled when either of the two predicates (EAT(x) or DRINK(x)) or both are met. For example, the phrase “I have eaten three apples for breakfast” is interpretable because EAT (“I have eaten three apples for breakfast”) is true.

$$\text{EAT}(x) \rightarrow \exists v, fp \mid \text{EATING}(v) \wedge \text{EATER}(v, \text{SPEAKER}) \wedge \text{FOOD_INGESTED}(v, fp) \quad (2)$$

The predicate EAT(x) is decomposed in the rule (2) described below, where fp is the food-related phrase of the original sentence and v is the verb that indicates the eating or drinking action.

The predicate EAT(x) can be used to discern whether each syntagma obtained through syntactic parsing contains the necessary information to be interpreted as a valid sentence to make a food register. For the predicate EAT(x) (2) to be valid, there must be a syntagma that is cataloged as a verb within the “Actions” entities described in the previous section. The subject who will perform the action does not have to appear explicitly in the sentence since it is assumed that the subject will always be the speaker (SPEAKER) acting as the protagonist of the food record. For example:

$$\begin{aligned} \text{EAT}(\text{“I have eaten three apples for breakfast”}) \rightarrow \text{EATING}(\text{“have eaten”}) \wedge \\ \text{EATER}(\text{“have eaten”,SPEAKER}) \wedge \text{FOOD_INGESTED}(\text{“have eaten”, “three apples for} \\ \text{breakfast”}) \end{aligned}$$

Additionally, the predicate FOOD_INGESTED(v, fp) (3) must be fulfilled with the rest of syntagmas (fp) that have been extracted from the sentence that is being interpreted. It checks whether they contain the food that must be registered and at the same time evaluates if the verb indicates the time of the intake (e.g. “I dined pasta yesterday”) or if there is a syntagma that expresses the temporal moment (e.g., “I had pasta for dinner yesterday”). The variable t is a chunk in the original sentence that specifies the timing (e.g., “for dinner yesterday”).

$$\text{FOOD_INGESTED}(v, fp) \rightarrow \exists fc, t \mid \text{FOOD_CHUNK}(v, fc) \wedge (\text{FOODHOUR}(v) \vee \text{FOODHOUR}(t)) \quad (3)$$

The predicate FOOD_CHUNK(v, fc) (4) evaluates the remaining phrases to extract the foods that will be registered, where fc is the food chunk that contains all the meals that are going to be recorded at the end of the process.

$$\text{FOOD_CHUNK}(v,fc) \rightarrow \exists f \mid \text{FOOD}(f) \wedge \text{EATED}(v,f) \wedge \text{Q_FOOD}(fc) \quad (4)$$

For example:

$$\text{FOOD_CHUNK}(\text{“have eaten”, “three apples”}) \rightarrow \text{FOOD}(\text{“apple”}) \wedge \text{EATED}(\text{“have eaten”, “apple”}) \wedge \text{Q_FOOD}(\text{“three apples”})$$

For FOOD_CHUNK (v, fc) to be true, there must be syntagmas (f) that correspond to a meal listed in the entities as “simple food or ingredients” (e.g., a tomato) or “compound foods or recipes” (e.g., a salad).

Additionally, the corresponding syntagmas must fulfill the predicate Q_FOOD(fc) (5) that analyzes the complements necessary to quantify the amount of food ingested. There are two different ways to express quantification, a direct number (e.g., “two apples”) or with containers or sizes (e.g., “a bowl of salad”, “a slice of pizza”). When containers and sizes are used, they can be also quantified (e.g., “two cans of tuna”, “a big piece of cake”). A direct number of aliments can only be used if they are countable.

In (5) nu indicates quantity, f represents the actual aliment in the food chunk, c is a container (e.g., a bowl) or indication of volume or size (e.g., a slice) and qc is a quantifier (e.g., two bowls) or size (e.g., big bowl) of the container.

$$\text{Q_FOOD}(fc) \rightarrow \exists nu, f \mid \text{QUANTIFIER}(nu) \wedge \text{COUNTABLE}(f)$$

OR

$$\text{Q_FOOD}(fc) \rightarrow \exists c, nu \mid \text{CONTAINER}(c) \wedge \text{QUANTIFIER}(nu) \wedge \text{N_CONTAINER}(c,nu) \quad (5)$$

OR

$$\text{Q_FOOD}(fc) \rightarrow \exists c, nu, qc \mid \text{CONTAINER}(c) \wedge \text{QUANTIFIER}(nu) \wedge \text{N_CONTAINER}(c,nu) \wedge \text{SIZE}(qc) \wedge \text{S_CONTAINER}(c,qc)$$

For example:

$$\text{Q_FOOD}(\text{“three apples”}) \rightarrow (\text{QUANTIFIER}(\text{“three”}) \wedge \text{COUNTABLE}(\text{“apple”}))$$

$$\text{Q_FOOD}(\text{“a bowl of cereals”}) \rightarrow (\text{CONTAINER}(\text{“bowl”}) \wedge \text{QUANTIFIER}(\text{“a”})) \wedge$$

$$N_CONTAINER(“bowl”, “a”)$$

$$Q_FOOD(“two big pieces of cake”) \rightarrow (CONTAINER(“pieces”) \wedge QUANTIFIER(“two”) \wedge N_CONTAINER(“pieces”, “two”) \wedge SIZE(“big”) \wedge S_CONTAINER(“pieces”, “big”))$$

A similar process is followed with beverages (equations 8 to 11) following the same process as with Equation (1).

$$DRINK(x) \rightarrow \exists v \mid DRIKING(v) \wedge DRINKER(v, SPEAKER) \vee DRINK_INGESTED(dp) \quad (6)$$

$$DRINK_INGESTED(v, dp) \rightarrow \exists dc, t \mid DRINK_CHUNK(v, dc) \wedge (FOODHOUR(v) \vee FOODHOUR(t)) \quad (7)$$

$$DRINK_CHUNK(v, dc) \rightarrow \exists b \mid BERVAGE(b) \wedge DRANK(v, b) \wedge Q_DRINK(bc) \quad (8)$$

$$Q_DRINK(bc) \rightarrow \exists c, nu, qc \mid CONTAINER(c) \wedge QUANTIFIER(qc) \wedge N_CONTAINER(c, nu) \wedge SIZE(qc) \wedge S_CONTAINER(c, qc) \quad (9)$$

After completing this process, the meals and complements to be registered have been identified. If during the process they do not fulfill some of the predicates, the missing pieces of information are identified so that the nutritional coach can request them to the user engaging in a conversation. If the failing parameters cannot be identified, then the phrase is interpreted as unknown and the coach can ask the user to rephrase.

4. Validation and Results

To evaluate our proposal, we have implemented it in Java using the Stanford-core-NLP parser [18] for the syntactic analysis. We have compared the accuracy of our proposal with respect to the keyword-spotting mechanism that is the standard used in current nutritional coaches. To do it, we have used a battery of 50 sentences in natural language (English).

For evaluation we have used simulated databases that replace the ones that should be used in a real implementation in order to have a single nutritional information result to be used for comparison. Once each sentence was interpreted by each procedure, the corresponding nutritional information was recorded and compared to human interpretation. Human interpretation was carried out computing the nutritional information corresponding to the food, quantifiers and other relevant information that was manually identified.

We define an interpretation as successful if the nutritional information computed agrees with the human annotation. There will be an error when the calculation does not exactly match the expected one.

The phrases that have been used for the validation have been designed to encode all the different structures, verbal actions and variability in the number of arguments (meals) to be recorded. Consequently, there are sentences that only describe the registration of one single food (e.g., “I have eaten one apple”), several foods (e.g., “I have had one apple, a zero coke and two slices of margherita pizza”), specific foods with descriptions of how they are composed (e.g., “I have eaten an omelet with two eggs and with two slices of cheese”), or a mixture (e.g., “for lunch I have had two apples and an omelet with two eggs and two slices of cheese”). The general structure used are sentences where the subject may be missing, and there is a verb as well as a variable amount of predicates that will be treated as the arguments to register in the nutritional record.

The results obtained are shown in Table 1. As can be observed, our proposal achieves 96% accuracy, while keyword-spotting obtains 44% accuracy.

Table 1. Results of the validation process by sentence type in the nutritional context.

Type of Sentence	Number of Sentences	Results			
		Semantic Parsing Procedure (Our Proposal)		Keywords Parsing (Dialog Flow ...)	
		Number of Correctly Interpreted Sentences	Number of Sentences with Interpretation Errors	Number of Correctly Interpreted Sentences	Number of Sentences with Interpretation Errors
Single food (1 argument)	13	13	0	13	0
More than one meal (1+ arguments)	12	12	0	9	3
Single Compound meal (1 argument)	13	13	0	0	13
More than one meal with compounded meals (1+ arguments)	12	10	2	0	12

Although both procedures have a similar accuracy when analyzing sentences with simple foods, when analyzing sentences with complex foods our proposal clearly outperforms the keyword approach. This is because the keyword analysis does not establish the dependencies between the different elements, and therefore the process of identifying food descriptions with a variable structure is not effective.

Regarding the nature of the errors, our approach only made mistakes in the interpretation of sentences that involved both complex foods and other foods due to the syntactic ambiguities in the first analysis process. These ambiguities are inherent to natural language. In many occasions when performing the syntactic analysis of sentences there may have different syntactic interpretations, without further information there are mismatches between the output produced and the expected interpretation. For example, “I have eaten a dish of meat with potatoes” can be interpreted as having eaten meat and potatoes on the side or that it was a recipe that includes meat and potatoes. To solve these problems, knowledge discovery approaches can be applied to the recipes databases that define which are the most probable constructions.

On the other hand, the approach based on keywords, failed in all cases in which the identification of the dependencies in the complex foods could not be made and thus the relations between the meals and their quantifiers and qualifiers is not accurately established.

5. Conclusions

Virtual nutritional coaches help professional nutritionists to guide and teach users the importance of healthy habits. For this reason, the inclusion of new ways of capturing such habits is an important step further in the health field. Current approaches are mainly based on keyword spotting and lack flexibility, diminishing the possibilities for detailed food records. In this paper, we have proposed a procedure for natural language understanding in the virtual nutritional coaching field based on syntactic and semantic analysis. This mechanism aims to improve the food recording process by giving more flexibility in the linguistic structures allowed in the system. We have defined entities and semantic rules to validate the syntactic structures and designed a procedure to automatically use semantic rules for daily food registration using natural language. We have designed an algorithm with which to perform automatic calculations of nutrients in complex foods using the proposed procedure, which allows to identify intake schedules, quantifiers, containers and the ingredients that make up the meals. The main difference between this mechanism and the processing of keywords is that with keyword processing it is not possible to know which keywords or tokens accompany or complement other keywords in the same sentence, losing valuable information.

We have implemented a version of the proposed procedure and validated it against human interpretation, comparing the results to the ones obtained by state-of-the-art keyword spotting

approaches. The results show that our proposal is more effective and obtains a higher accuracy rate even with complex sentences.

With these advances it is possible to build nutritional coaches with improved natural language understanding capabilities, giving the possibility to the users to describe the ingestions more precisely and obtaining more accurate and detailed nutritional interpretations for the system to provide a more accurate coaching. For future work, we intend to implement a dialogue manager that makes it possible to ask the user for relevant missing elements when the interpretation cannot be completed.

Acknowledgments: This research has been supported by the project DEP2015-70980-R of the Spanish Ministry of Economy and Competitiveness (MINECO) and European Regional Development Fund (ERDF), as well as, received inputs from the COST Action IC1303 AAPELE.

References

1. Azar, K.M.; Lesser, L.I.; Laing, B.Y.; Stephens, J.; Aurora, M.S.; Burke, L.E.; Palaniappan, L.P. Mobile Applications for Weight Management: Theory-Based Content Analysis. *Am. J. Prev. Med.* **2013**, *45*, 583–589.
2. Ruiz-Zafra, Á.; Benghazi, K.; Noguera, M.; Garrido, J.L. Zappa: An open mobile platform to build cloud-based m-health systems. In *Ambient Intelligence-Software and Applications*; Springer: Heidelberg, Germany, 2013; pp. 87–94.
3. McTear, M.; Callejas, Z.; Griol, D. *The Conversational Interface—Talking to Smart Devices*; Michael McTear; Springer: Heidelberg, Germany, 2016; Available online: <https://www.springer.com/gp/book/9783319329659> (accessed on 9 May 2018).
4. Potter, D. W., Powell, K. R., Humphreys, K. W., & Hamilton, J. S. (2008). U.S. Patent No. 7,409,337. Washington, DC: U.S. Patent and Trademark Office.
5. Simpson, N.D. System and Method for Controlling the Operation of a Device by Voice Commands. 24-Jul-2008. Available online: <https://patentimages.storage.googleapis.com/51/26/62/ee24e38198fef7/US7774204.pdf> (accessed on 2 May 2018).
6. Natural Language Toolkit—NLTK 3.3 Documentation. Available online: <https://www.nltk.org/> (accessed on 7 May 2018).
7. Manning, C.D.; Manning, C.D.; Schütze, H. *Foundations of Statistical Natural Language Processing*; MIT Press: Cambridge, MA, USA, 1999.
8. Jurafsky, D.; Martin, J.H. *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition*; Prentice Hall: Upper Saddle River, NJ, USA, 2009.
9. Tur, G. *Spoken Language Understanding: Systems for Extracting Semantic Information from Speech*; Microsoft Research; John Wiley & Sons: Hoboken, NJ, USA, 2011.
10. Huffman, M. Health Coaching: A New and Exciting Technique to Enhance Patient Self-Management and Improve Outcomes. *Home Healthc. Now* **2007**, *25*, 271.
11. Olsen, J.M. Health Coaching: A Concept Analysis. *Nurs. Forum* **2014**, *49*, 18–29.
12. Bickmore, T.; Gruber, A.; Picard, R. Establishing the computer–patient working alliance in automated health behavior change interventions. *Patient Educ. Counsel.* **2005**, *59*, 21–30.
13. Mansfield, E.A.; Kocher, J.P. Device and Method for Monitoring Dietary Intake of Calories And Nutrients. 15-Aug-1996. Available online: <https://patents.google.com/patent/US5819735A/en> (accessed on 8 May 2018).
14. Stephens, J.; Allen, J. Mobile Phone Interventions to Increase Physical Activity and Reduce Weight. *J. Cardiovasc. Nurs.* **2013**, *28*, 320–329.
15. Ochoa, S.F.; Gutierrez, F.J. Architecting E-Coaching Systems: A First Step for Dealing with Their Intrinsic Design Complexity. *Computer* **2018**, *51*, 16–23.
16. Dialogflow. Dialogflow. Available online: <https://dialogflow.com/> (accessed on 4 May 2018).
17. Popescu, A.M.; Armanasu, A.; Etzioni, O.; Ko, D.; Yates, A. Modern natural language interfaces to databases: Composing statistical parsing with semantic tractability. In Proceedings of the 20th international conference on Computational Linguistics, Geneva, Switzerland, 23–27 August 2004; p. 141.
18. Liang, P. Lambda Dependency-Based Compositional Semantics. *arXiv* **2013**, arXiv:1309.4408.

19. Wang, Y.; Berant, J.; Liang, P. Building a Semantic Parser Overnight. In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*; Association for Computational Linguistics: Beijing, China, 2015; pp. 1332–1342.
20. Liang, P. Learning executable semantic parsers for natural language understanding. *Commun. ACM* **2016**, *59*, 68–76.
21. Apache OpenNLP. Available online: <https://opennlp.apache.org/> (accessed on 7 May 2018).
22. Software—The Stanford Natural Language Processing Group. Available online: <https://nlp.stanford.edu/software/> (accessed on 6 May 2018).



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).