ISCA Archive
http://www.isca-speech.org/archive

SPECOM'2004: 9th Conference
Speech and Computer
St. Petersburg, Russia
September 20-22, 2004

Ontology of the subject domain
"Speech signals recognition and synthesis".

Valery I. Galunov (1), Boris M. Lobanov (2) and Nikolay G. Zagoruiko (3),

(1) St. Petersburg State University, 7/9, Universitetskaya emb., St. Petersburg, 199034, Russia
auditech@online.ru
(2) United Institute of Informatics Problems NASB, Surganov Str., 6, Minsk 220012, Belarus
lobanov@newman.bas-net.by
(3) Institute of Mathematics SD RAS, pr. Koptyug 4, Novosibirsk, 630090, Russia
zag@math.nsc.ru

Abstract

The general structure of ontology of subject domain is described. Definitions of basic elements of ontology are given. Examples of the description a some elements of ontology of the subject domain "Speech signals recognition and synthesis" are resulted. The possible variant of the organization of collective development of such ontology is offered.

1. Introduction

Ontology (O) is the brief description of a subject domain structure, which includes terms (T), the designating allocated fragments of a subject domain (objects, concepts, the phenomena, attributes and so forth), relations (R) between them, and definitions (D) these terms and relations [1,2]:

$$O = < T, R, D >.$$

In graphic performance the ontology can look like a pyramidal network, which tops are designated by terms of a subject domain, and edges specify relations between these terms. The root top contains the term, which is the name of a subject domain. Tops of the second level designate basic elements of the description of a subject domain or base categories. They are connected to root top the relation "part – whole". Each category is connected to tops of the following level (with concepts) relations "general – private". Concepts are in turn described through other more simple concepts. Between any pair tops of a network the appropriate relations can be specified: "it is used for", "contains in", "is a part" etc.

It is supposed, that ontology it will be connected to system of automatic extraction of knowledge from texts in natural language which will accumulate and update the knowledge base of the given subject domain. The ontology will play a role of the basic directory with which help this system will determine, whether the document is connected to a subject domain and to what category or to what concepts concern the document as a whole or its separate semantic fragments.

Ontology is called to structure and order knowledge and also to unify a terminology of the given subject domain, that it is useful for the following purposes:
- Perfection of the organization of researches in the given subject domain;
- Improvement of methods of learning;
- Improvement of quality of search engines.

Ontology can become claimed only in the event that it grows out collective work of the broad audience of experts at given subject domain.

2. Nucleus of ontology

Let's explain in the beginning, that such a subject domain. Such sphere of activity of people in which are considered the set of material or ideal fragments allocated with any purpose (objects, concepts, properties, problems, processes, the phenomena, relations etc.) refers to as a subject domain. The given variant of ontology is focused on support of creation ontologies mainly such subject domains which are connected to the analysis, synthesis and transformation of the information about any fragments of the real world (measurement, accumulation of the data, discovering of regularities (knowledge), storage, processing and data and knowledge transfer, use of knowledge for recognition, forecasting and synthesis).

Now about structure of a nucleus includes such elements which are used in ontologies any subject domain.

2.1. Terms (T)

The symbolical designation of any allocated fragment of a subject domain (object, concept, process and so forth) refers to as the Term. Examples of terms: water, training, rent etc. Values of terms are their characteristics at a semantic level and are defined by those functions ("to satisfy thirst", "to extinguish a flame", "to transfer knowledge", "temporarily to use"). Each value of the term is defined by a set of characteristics the appropriate fragments of the described world.

In considered variant of ontology the set of Terms consists of set of the ***Categories*** consisting of sets ***of***

Concepts. In language OWL the Term is *an abstract class*. The category is the *concrete class* being the descendant of the Term. The category consists of set *of values - objects* (Concepts).

At definition of the term it is desirable to result the list of other terms, which meet in the literature according to a subject domain and are synonyms of the determined term.
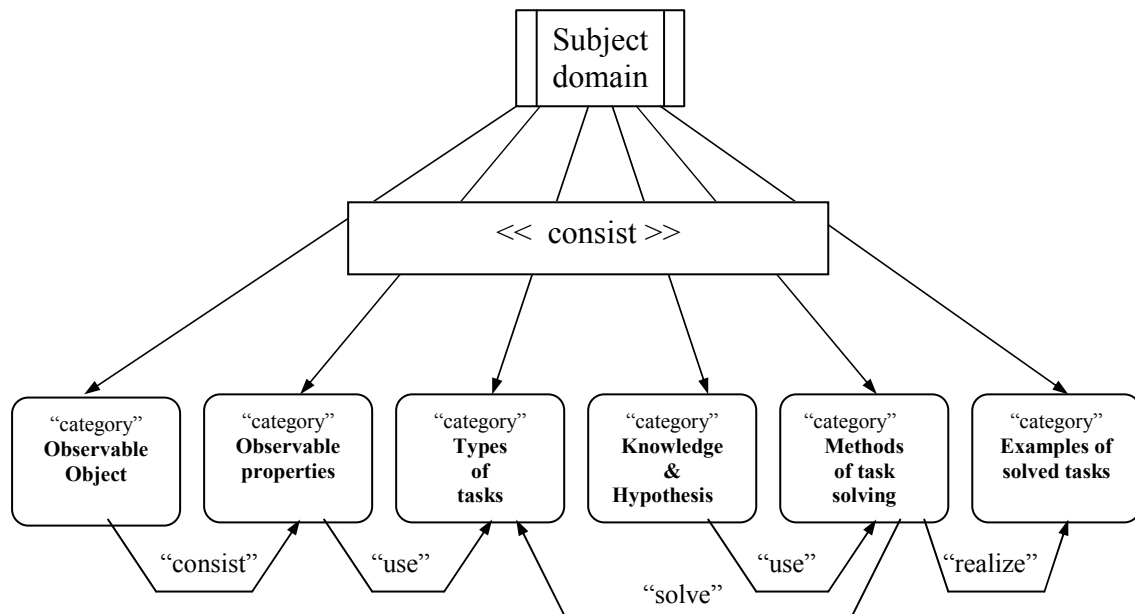
```
                    ┌─────────────┐
                    │   Subject   │
                    │   domain    │
                    └─────────────┘
          ┌─────────────────────────────────┐
          │         << consist >>           │
          └─────────────────────────────────┘
```

| "category" Observable Object | "category" Observable properties | "category" Types of tasks | "category" Knowledge & Hypothesis | "category" Methods of task solving | "category" Examples of solved tasks |

"consist"   "use"   "solve"   "use"   "realize"

Figure 1. Basic Categories of ontology

2.2. Basic categories

Terms, which are used in ontologies **various subject domains** refer to as basic categories. If subject domains are connected to reception, the analysis, processing, storage and data and knowledge transfer the common categories for them are the following:
- *Observable Objects;*
- *Observably properties;*
- *Types of tasks;*
- *Knowledge and a hypothesis;*
- *Methods of the task solving;*
- *Examples of solved tasks.*

2.2.1. Observable Objects
The given category represents observably and investigated objects of a subject domain. That allocated fragment of a subject domain (object, characteristic, process, phenomenon and so forth), which properties are measured and fixed in reports of supervision and then used at the decision of various tasks of the given subject domain. Examples of observably objects can be "patients", "earthquakes", " speech signals ", " weather conditions" etc.

2.2.2. Observable properties
This category represents in ontology properties (characteristics), which are observed and studied at observably objects of a subject domain. Represents means and methods for definition of values measuring characteristics, and means of display of the received results in symbolical or graphic forms. At a mention of widely known devices or methods it is possible to be limited to their names (for example, "microphone", "voltmeter", "oscillograms"). In other cases descriptions and instructions of a various degree of a detail can be demanded.

2.2.3. Types of Tasks
The given category unites the general problems (tasks), which reflect the purposes and sub purposes of the activity of each subject domain. They should be described with that degree of a detail which is necessary for their unequivocal understanding experts of the given area.

Various tasks differ from each other by the following elements:
$Z = < G, O, D, S, F >$. Here:

G - the description of the purpose in terms of desirable values of the given characteristics;
O - means of supervision over the current values of these characteristics;
D - means of influence for values of characteristics;
S - ways of a choice of the necessary means of influences;
F - criteria for definition of the moment of achievement of the purpose.

2.2.4. Knowledge and a hypothesis

In this section are briefly described knowledge and hypothesis which are used at the description and the solving of tasks of the given subject domain. The description should be accompanied by links to sources, which contain the detailed information on the mentioned questions.

2.2.5. Methods of Tasks Solving

Here the methods used for the solving of various tasks of the given area are described. For each method and the appropriate tasks the relation such as « is underlined is applied for », « is used at » etc. links to primary sources with the detailed description of methods are resulted.

2.2.6. Examples of solved tasks.

Examples of tasks of various type which were solved or are solved now are resulted. The publications containing more detailed information on these tasks are specified.

2.2. Concepts

At a hierarchical level of ontology, the following a level of Categories, there are terms which designate semantic units (Concept) of the given subject domain. In each subject domain the specific Concepts are used. In ontological graph the concepts *Category III.* are nods connected to nods of Categories by relations of a class "private - general". Concepts can be defined or direct in natural language, or with the help of finnier Concepts. At the description of Concepts of various levels it is necessary to specify the appropriate relations between them.

3. Relations (R)

The kind of connection between terms refers to as the relation, character of dependence, a kind of influence, the fact of applicability etc. In the field of information technologies the greatest distribution have relations of two classes: "Aggregation" and "Use".

3.1. Relate

It is a base name of relation for all kind of relations between concepts.

3.2. Aggregate

It relation expresses semantics of a situation in which one concept aggregates other concept. The group "aggregation" includes the relations designating connections such as "consist", "private – general", "belongs", "part of" etc.

3.3. Use

The relation "Use" expresses semantics of a situation in which one Concept uses other Concept. The group "Use" includes the relations designating connections such as "is used for", "is applicable to", "is intended for" etc.

3.4. Other relations

In different cases can be demanded and other sets of relations: mention, solve, realize etc.

4. Ontology of the subject domain "Speech signals recognition and synthesis"

Using the structure of ontology described above we shall result brief sketches of its possible filling for a subject domain connected to recognition and synthesis of speech signals [3]. The maintenance of each base category is described by a set *of concepts.* We shall result the list of these concepts.

Concepts of Category 1 "Observable Objects"

1.1. Speech signals
1.2. Speakers
1.3. Emotions and a status of speaker
1.4. Operational environment

Concepts of Category 2 "Observably properties"

2.1. Receivers of a sound
2.2. Means of preliminary processing
2.3. Selected characteristics
2.4. Ways of graphic performance
2.5. Protocols of supervision

Concepts of Category 3. Types of solving tasks

3.1. Systems of speech signals recognition
   3.1.1. Systems of recognition and understanding of sense told
   3.1.2. Systems of speaker verification and identification
   3.1.3. Systems of recognition emotional and physiological status of the speaker
3.2 Systems of speech verification and speech identification.
3.3. Systems of speech signals synthesis

3.2.1. Systems of synthesis of speech under the text
3.2.2. Systems of speaker voice cloning

Concepts of Category 4. ***Knowledge and hypothesis.***

4.1. Models of human speech production and perception systems
4.2. Models of the speech communications
4.3. Physiology of hearing system
4.4. Psychoacoustics
4.5. Phonetics, morphology, syntax, semantics
4.6. Measures of affinity in space of speech signal characteristics
4.7. Rules and criteria of acceptance of decisions at recognition
4.8. Ergonomics of speech systems
4.9. Interrelation between all area of scientific knowledge and hypothesis.

Concepts of Category 5. ***Methods of the task solving***

5.1. Mathematical methods of pattern recognition
5.1.1. Methods of decision-making
5.1.2. Methods of a choice of the informative
   features
5.1.3. Methods of automatic classification
5.2. Methods of measurement of distances between objects
5.2.1. Types of measuring scales
5.2.2. A measure of affinity between objects
5.3. Methods of the analysis of speech signals
5.3.1. The spectral analysis
5.3.2. The cepstral analysis
5.3.3. The wavelet analysis
5.3.4. The fractal analysis
5.3.5. On the foundation of models of speech perception module.
5.4. Methods of speech signals synthesis
5.4.1. Compilative synthesis
5.4.2. Formant synthesis
5.4.3. Synthesis on speech production models
5.5. Methods of understanding of speech messages
5.5.1. Methods of parse
5.5.2. Methods of the semantic analysis
5.5.3. Methods of the pragmatically analysis
5.5.4. Methods of the organization of speech dialogue

Concept of Category 6. **Examples of the solved**

6.1. Systems of recognition

6.1.1. Recognition of isolated commands
6.1.2. Recognition of conjoint speech
6.1.3. Recognition of speakers
6.1.4. Recognition of an emotional status of the pilot
6.2. Systems of synthesis
6.2.1. Compilative synthesis
6.2.2. Formant synthesis
6.2.3. Synthesis on human speech tract models
6.2.4. Voice clones

5. Definitions of concepts

Let's result brief definitions of the concepts included in some base categories. To facilitate work of the future programs of automatic extraction of the semantic information from texts in natural language possible displays of each concept of text documents are resulted. In square brackets links to the literature are given.

Category 1. **Observably objects**

Concept 1.1. **Speech signals** - fluctuations of sound pressure on an input of the receiver, produced by natural or artificial speech production mechanisms. Depending on duration of an observably site of a speech signal tasks of the analysis and synthesis of speech segments, phonemes, syllables, words or word collocations are put. Signals of singing and whisper speeches are studied.
Displays in the text: a speech signal, sounds of speech, a speech material, sound types, a phoneme, a syllable, a word, a soundtrack, singing, whisper, …

Concept 1.2. **Speakers -** are people which speech signal is required to analyze (to recognize, understand, verify etc.) or to synthesize. The following characteristics concerning speakers have value: the sex, age, one or is a lot of the speakers, trained or untrained, whether presence of accents, quality of diction, defects of speech etc.
Displays in the text: the speaker, announcer, the operator speaking, a female voice, a man's voice, a children's voice, one speaker, any speaker, the trained announcer, the native speaker, accent, diction, normal speech, speech with defects, …

Category 3. **Types of solving tasks**

Concept 3.1. **Systems of recognition of speech signals**
3.1.1. Systems of recognition and understanding of sense told.
a) Systems of recognition of a small number separately said words (from 1 till 30-40 words) in which integrated characteristics of words (duration, average loudness, an average spectrum etc.) are used.
b) Systems of recognition of dictionaries of average volume (tens and hundreds words) in which characteristics of separate parts of a word are used: segments, sound types, syllables.
c) Systems of recognition of the given words in a stream of speech.

Concrete systems differ from each other required reliability, an opportunity of adaptation to speaker and to new dictionary, characteristics of an operational environment etc.

d) Systems of recognition of phonemes of conjoint speech.

Intend for recognition of phonemes in a stream of conjoint speech at the unlimited dictionary (automatic shorthand).

e) System of understanding of speech messages.

Systems with different volume of the dictionary in which for increase of recognition reliability it is used aprioristic knowledge not only of phonetics, lexicon and syntax, but also of semantics and the pragmatist of speech acts.

Displays in the text: recognition of the commands, isolated words, conjoint speech, speech management, automatic shorthand, volume of the dictionary, reliability, adaptation to speaker …

Concept 3.2 **Systems of speech verification and speech identification.**

3.2.1. Systems of speaker verification.

3.2.2. Systems of speaker identification.

Displays in the text: speeker verification, identification, individual speech characteristics,…

Concept 3.3. **Systems of synthesis of speech signals**

3.2.1. Systems of synthesis of speech under the text.

Intend for automatic synthesis of the speech message under the text. The basic requirements - high legibility and naturality a speech signal. Circuits of transitions "the text - a phonetic transcription - commands on elements of a sound wave generation" are used. As elements of generation parameters of speech production model or short pieces of a speech signals can serve.

3.2.2. Systems of cloning voices of the speaker.

These systems should reproduce specific features of speech of the concrete person (height of the basic tone, a timbre, rate, diction etc.). Thus it is possible to use or short pieces of a speech signal of this person, or exact his model of speech path (tract?).

Displays in the text: synthesis of speech, a synthesizer, speaking machine, microwave synthesis, compilative synthesis, model of speech production, an intonational contour, ...

Category 5. **Methods of the task solving**

Concept 5.1. **Mathematical methods of pattern recognition**

5.1.1. Methods of construction of decision functions
Statistical deciding rules. Logic deciding rules.Algebraic deciding rules. Heuristic deciding rules.

5.1.2. Methods of a choice of the informative features.

Method of random search with adaptation. Directed taxonomical search of attributes. Relaxation method (algorithm AdDel).

5.1.3. Methods of automatic classification
Algorithms of taxonomy. Algorithms of class FOREL and KRAB.

Displays in the text: deciding rules, decision function, training to recognition, a choice of attributes, taxonomy...

6. How to organize development of the ontology?

Authors do not apply for indisputability of the performances about ontology of such complex scientific direction what the area of recognition and synthesis of speech is. The fragment of ontology described above is only its draft sketch. The ontology only then it will be claimed, if it will be result of collective work of the broad audience of experts of the given subject domain. The given text should be considered as invitations to the experts investigating the speech communications to take part in collective development of ontology our subject domain.

Elements of all ontology levels since the uppermost are subject to development and processing. To determine on which bottom level it is necessary to stop, we should take into account, that ontology should not be the full detailed directory or the encyclopedia on all questions of a subject domain. It should reflect basic elements give general performance about the device of the given subject domain, and to specify sources which contain more detailed information.

The organization of this collective work can be, for example, such: initiative working group (WG) of the qualified experts representing different parts of solved problems is created. This group on public principles (the sponsor or the external customer will not appear yet) undertakes to carry out next functions:

1. To formulate the problems demanding the collective decision.

2. To dispatch these formulations by e-mail to all wishing to participate in development of ontology (to "collective of developers").

3. To collect and process the arrived variants of the decision of problems. To develop variants of the decision, as much as possible taking into account arrived offers.

4. To dispatch these variants to members of collective of developers for discussion.

5. To make final variant of the decision of problems on the basis of the judgments stated by members of collective.

Works are now conducted directed on creation of programs, which should help to experts to create and accompany with machine versions of ontology subject domains. We hope, that collective efforts of speech community will allow creating one of the first ontology of the complex scientific area.

7. References

[1] Kleschev A.S., Artemieva I.K. Mathematical models of ontology of subject domains. Ch.1. Existing approaches to definition of concept "ontology". Scientific-Technical Information, Series 2. №2, 2001, pp.20-26. (in Russian)

[2] Gruber T.R. Toward Principles for the Design of Ontology Used for Knowledge Sharing // Int. J. of Human-Computer Studies.-1994.-Vol. 43 (5/6)-P.907-928.

[3] Survey of the state of the art in human language technology Ed. Chif Roland Cole, Cambridge University Press, 1997.