Automatic Linguistic Report of Traffic Evolution in Roads

Alberto Alvarez-Alvarez¹
Daniel Sanchez-Valdes¹
Gracian Trivino¹
Ángel Sánchez²
Pedro D. Suárez²

Abstract
In the field of Intelligent Transportation Systems, one important challenge consists of maintaining updated the electronic panels installed in roads with relevant information expressed in natural language. Currently, these messages are produced by human experts. However, the amount of data to analyze in real time and the number of available experts are imbalanced and new computational tools are required to assist them in this work. Moreover, the same problem appears when we deal with automatically generating linguistic reports to assist traffic managers that must take their decisions based on large amounts of quickly evolving information.

In this paper, we contribute to solve this problem by designing a computational application based on our research in the field of Computational Theory of Perceptions. Here, we present an application where we generate linguistic descriptions of the traffic behavior evolving in time and changing between different levels of service. We include some results obtained with both, simulated and real data.

Key words: Traffic Analysis, Linguistic Summarization, Computing with Perceptions

1. Introduction

Intelligent Transportation Systems (ITS) aim to get safer traffic conditions and comfort in transportation, and also to increase the road traffic efficiency by improving the functionality of cars and roads [8, 7, 5].

Due to increasing social demands of mobility and safety in road transportation and the increasing computer capabilities, the need of automatic, economic and real-time solutions for reliable traffic flow analysis becomes a priority for many governments. In this context, one goal of automatic traffic analysis is the detection and tracking of vehicles driving through a controlled area in order to discover abnormal events such as traffic congestions, speed violations, some other illegal behavior of drivers or even the detection of accidents.
The availability of new suitable computational applications certainly will improve the efficiency of roads, assisting in the quick detection of traffic alarms, and also helping to foresee some problems when traffic is normal in road and highways.

An interesting and paradigmatic problem consists of generating dynamically the most adequate natural language (NL) messages to communicate with drivers using electronic panels installed in the roads. Currently, these messages are produced by human experts but this task can be tough and tedious. Moreover, the balance between the amount of changing data to analyze and the number of experts available is getting worse dramatically. This situation causes the need of computational systems that can interpret and describe linguistically the large amount of available information.

In [11], we can found a survey of technologies for locating the position of vehicles on the road. In this direction, the works by W. Wen show an intelligent traffic management expert system with Radio Frequency Identification (RFID) technology [26] and a dynamic and automatic traffic light control expert system for solving road congestion problems [25]. In [19], authors present a technology to collect and organize data about the vehicles moving in a road network. Nevertheless, to the best of our knowledge, currently, a technology able to generate relevant linguistic descriptions of the traffic behavior is not available.

In this paper, we aim to contribute to this field by presenting a computational application able to generate linguistic descriptions in real-time about the traffic evolution. Our approach is based on the use of Fuzzy Logic (FL), which is widely recognized for its ability for linguistic concept modeling and its use in system identification [21]. On the one hand, semantic expressiveness, using linguistic variables [31, 32, 33] and rules [15, 30], is quite close to NL. On the other hand, being universal approximators [9] fuzzy inference systems are able to perform nonlinear mappings between inputs and outputs. More specifically, our approach is based on the Computational Theory of Perceptions (CTP) introduced in the Zadeh’s seminal paper “From computing with numbers to computing with words - from manipulation of measurements to manipulation of perceptions” [29] and further developed in subsequent papers. CTP provides a framework to develop computational systems with the capability 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 on this line, we have generated linguistic descriptions of different types of phenomena. For example, we generated financial reports from data taken from the Spanish Securities Market Commission (CNMV) [18] and linguistic descriptions about relevant features of the Mars’ Surface [2]. Specifically, in the field of ITS, we generated linguistic reports about the traffic on roundabouts [23] and we generated assessing reports in truck driving simulators [13, 14].

In this work, we focused on the perception of change. We explored possibilities to perform linguistic descriptions of how the traffic evolves in time. We have researched on how to model the meaning of sentences such as “the phenomenon is changing from state A to state B”. In order to model the evolution of phenomena in time, we have used our previous works on Fuzzy Finite State Machines (FFSMs). Here, we have extended the use of the FFSM’s output function to be used with this aim. With a different approach, see in [20] how this idea has also been explored with the aim of summarizing network flow statistics.

The remainder of this paper is organized as follows. Section 2 presents the main concepts of our approach to linguistic description of complex phenomena evolving in time. Section 3 describes how to use these concepts for the linguistic description of the traffic behavior. Afterwards, Section 4 describes the experimentation carried out. Finally, Section 5 draws some conclusions and introduces some future research works.

2. Linguistic description of complex phenomena

In this section, we present several basic concepts of our contribution to CTP aimed to develop computational systems able to generate linguistic descriptions of phenomena. According to Zadeh, the object of perceptions are not only the attributes of objects, e.g., the distance, velocity and angle. The object of perceptions can be the whole systems, e.g., a person parking a car, the traffic in a roundabout, the air-conditioned system in buildings, etc. In this way, we use the term phenomenon to represent an object, or a set of interrelated objects, that is perceived in the computer environment. Phenomena are located in certain context and evolve in time among different situation types.
The Granular Linguistic Model of a Phenomenon (GLMP) is based on subjective perceptions of a domain expert that we call designer. The more experienced designer, with better understanding and use of NL, the richer the model with more possibilities of achieving and responding to final users’ needs and expectations. The designer uses the resources of the computer, e.g., sensors, to acquire data about a phenomenon and uses her/his own experience to interpret these data and to create the model. Then, the designer uses the resources of the computer to produce the linguistic utterances. In the following subsections, we introduce the main elements of our architecture for the linguistic description of complex phenomena.

2.1. Computational Perception (CP)

A CP is the computational model of a unit of information acquired by the designer about the phenomenon to be modeled. In general, CPs correspond to particular details of the phenomenon at certain degrees of granularity. A CP is a couple \((A,W)\) where:

\[ A = (a_1,a_2,\ldots,a_n) \]

is a vector of \(n\) linguistic expressions (words or sentences in NL) that represents the whole linguistic domain of the CP. Each \(a_i\) describes the value of the CP in each situation with specific granularity degree. These sentences can be either simple, e.g., \(a_i = \text{"Traffic density is high"}\) or more complex, e.g., \(a_i = \text{"Usually, at midday, the traffic density increases in this part of the road"}\).

\[ W = (w_1,w_2,\ldots,w_n) \]

is a vector of validity degrees \(w_i \in [0,1]\) assigned to each \(a_i\) in the specific context. The concept of validity depends on the application, e.g., it is a function of the truthfulness and relevance of each sentence in its context of use.

In this application paper, in order to model our perception of temporal evolution of phenomena, we applied a paradigm composed of three types of CP, namely, the perception of the current state (assertive CP), the perception of the trend to evolve (derivative CP), and the summary of accumulated perceptions.
The assertive CP is associated with a linguistic expression of the current state of a characteristic of the phenomenon, e.g., “the traffic density is high”. The derivative CP corresponds to trend analysis information and gives insight into how the phenomenon is evolving in time, e.g., “the traffic density is decreasing”. Finally, the integrative CP represents the accumulated perception of the phenomenon over a period of time, e.g., “the traffic density in the last period has been low”.

2.2. Perception Mapping (PM)

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

\[
U = (u_1, u_2, \ldots, u_n)
\]

is a vector of \(n\) input CPs \(u_i = (A_{u_i}, W_{u_i})\). In the special case of first order perception mappings \((1-PMs)\), these are the inputs to the GLMP and they are values \(z \in \mathbb{R}\) being provided either by a sensor or obtained from a database.

\[
y = (A_y, W_y)
\]

is the output CP.

\[
W_y = g(W_{u_1}, W_{u_2}, \ldots, W_{u_n})
\]

is an aggregation function employed to calculate \(W_y = (w_1, w_2, \ldots, w_n)\) from the input CPs. In FL, many different types of aggregation functions have been developed. For example, \(g\) might be implemented using a set of fuzzy rules. In the case of 1-PMs, \(g\) is built using a set of membership functions \(\mu_{a_i}(z)\) as follows:

\[
W_y = (\mu_{a_1}(z), \mu_{a_2}(z), \ldots, \mu_{a_n}(z)) = (w_1, w_2, \ldots, w_n)
\]

\(T\) is a text generation algorithm that allows generating the sentences in \(A_y\). In simple cases, \(T\) is a linguistic template, e.g., “Road vehicle density is \{high | medium | low\}”.

There are many types of PMs. In this paper, we contribute to this research line by exploring two of them focused on describing how phenomena evolve on time. In Section 3.2.8, we introduce a PM based on fuzzy quantifiers and in Section 3.2.9, we introduce a PM based on a FFSM.

2.3. Granular Linguistic Model of a Phenomenon (GLMP)

The GLMP consists of a network of PMs. Each PM receives a set of input CPs and transmits upwards a 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 degree of granularity. Using different aggregation functions and different linguistic expressions, the GLMP paradigm allows the designer to model computationally her/his perceptions.

Fig. 1 shows an example of a GLMP. In this example, we describe the phenomenon at a very basic level in terms of three input variables that provide values \(z_1, z_2,\) and \(z_3\) respectively at a certain instant of time. These variables are introduced in the perception mappings \(PM_1, PM_2\) and \(PM_3\), providing \(CP_1, CP_2\) and \(CP_3\). Using these three 1-CPs, we use the perception mappings \(PM_4\) and \(PM_5\) to explain \(CP_4\) and \(CP_5\). Finally, a 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, 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.

2.4. Report generator

Fig. 2 shows the basic architecture of our report generator. The main processing modules of this computational system are, namely, the Data Acquisition (DAQ) Module, the Validity Module, and the Expression Module that are described in the following sections.

2.4.1. DAQ Module

This processing module provides the data needed to feed the 1-CPs. The Data Acquisition module provides the interface with the application physical environment. This module could include either sensors or access to information in a database. In this paper we generate examples of these data using both, a road traffic simulator and an image processing module.
2.4.2. Validity module

Once a sample of input data is available, the Validity module uses the aggregation functions in the GLMP to calculate the validity degree of each \( CP \). Therefore, this module provides as output a collection of linguistic clauses together with associated degrees of validity.

2.4.3. Expression module

Provided a set of valid linguistic clauses, the goal is to combine this information to build a linguistic report. This module deals with generating the most relevant linguistic report by choosing and connecting the adequate linguistic clauses based on a Report Template data structure.

3. Linguistic Description of Traffic Behavior

This section describes how to apply our approach for linguistic description of complex phenomena to the description of the traffic behavior. We explain the relevant modules needed to produce the linguistic description of traffic behavior.

From the study of several sources, including the Highway Capacity Manual [17], we obtained a first list of parameters about the traffic behavior that our reports should contain, namely, the speed of vehicles, traffic density and level of service in road (LOS). These parameters allow us to report the traffic behavior and to study anomalous situations that may occur, e.g., a vehicle traveling at a speed too high or too low, a vehicle circulating in opposite direction. To enrich the report, the linguistic description should include a time reference to place each event in the fraction of time in which it occurred.

3.1. DAQ Module

Here, we used basic measures of traffic parameters that can be obtained from different type of sensors, e.g., video cameras, radar, pressured hoses and inductive burial loops. These basic measures are, namely, the vehicles speed \( (vs) \), the average road speed \( (rs) \), that is calculated as the average of speeds at each moment, and the traffic density \( (td) \), which is calculated as the percentage of road that is occupied at each time instant.

3.2. Validity module

Fig. 3 shows a GLMP which tries to summarize and highlight the relevant aspects of the traffic behavior. In the following subsections, we describe each of the \( PMs \) and associated \( CPs \) in this model.
3.2.1. Road Speed (1-PM<sub>RS</sub>)

It is an assertive 1-PM whose input is the numerical value of the average road speed \((rs \in \mathbb{R})\). The output CP<sub>RS</sub> includes the following set of NL sentences:

\[ a_{RS1} \rightarrow \text{“The average road speed is very low”} \]
\[ a_{RS2} \rightarrow \text{“The average road speed is low”} \]
\[ a_{RS3} \rightarrow \text{“The average road speed is medium”} \]
\[ a_{RS4} \rightarrow \text{“The average road speed is high”} \]
\[ a_{RS5} \rightarrow \text{“The average road speed is very high”} \]

As in the rest of 1-PMs described in the following subsections, the validity degrees are obtained by means of a set of uniformly distributed trapezoidal membership functions (MFs) forming a strong fuzzy partition (SFP) [22]. Here, e.g., the validity degrees are directly \( w_{RS1} = V_{RS}(rs) \), \( w_{RS2} = L_{RS}(rs) \), \( w_{RS3} = M_{RS}(rs) \), \( w_{RS4} = H_{RS}(rs) \), and \( w_{RS5} = V_{HRS}(rs) \).

3.2.2. Road Speed Trend (1-PM<sub>RST</sub>)

This derivative 1-PM has validity degrees obtained from the numerical derivative of \( rs \) (\( drs/dt \)). It allows to generate the following set of NL sentences:

\[ a_{RST1} \rightarrow \text{“The average road speed is decreasing”} \]
\[ a_{RST2} \rightarrow \text{“The average road speed is keeping constant”} \]
\[ a_{RST3} \rightarrow \text{“The average road speed is increasing”} \]

3.2.3. Road Speed in the Last Period (1-PM<sub>RSLP</sub>)

It is an integrative 1-PM whose input is the numerical value of the average road speed \((rs \in \mathbb{R})\). The output CP<sub>RSLP</sub> includes the following set of NL sentences:

\[ a_{RSLP1} \rightarrow \text{“The average road speed in the last period was very low”} \]
The validity degrees are obtained by means of the aggregation function $g_{RSLP}$, which calculates the average value of $r_s$ during a period (75) defined empirically.

3.2.4. Traffic Density perception mapping (1-$PM_{TD}$)
It is an assertive 1-$PM$ that produces the following set of NL sentences:

- $a_{TD_1} \rightarrow \text{“The traffic density is extremely low”}$
- $a_{TD_2} \rightarrow \text{“The traffic density is very low”}$
- $a_{TD_3} \rightarrow \text{“The traffic density is low”}$
- $a_{TD_4} \rightarrow \text{“The traffic density is high”}$
- $a_{TD_5} \rightarrow \text{“The traffic density is very high”}$
- $a_{TD_6} \rightarrow \text{“The traffic density is extremely high”}$

3.2.5. Traffic Density Trend (1-$PM_{TD_T}$)
This derivative 1-$PM$ has validity degrees obtained from the numerical derivative of $td$ ($dtd/dt$). It allows to generate the following set of NL sentences:

- $a_{TD_T_1} \rightarrow \text{“The traffic density is decreasing”}$
- $a_{TD_T_2} \rightarrow \text{“The traffic density is keeping constant”}$
- $a_{TD_T_3} \rightarrow \text{“The traffic density is increasing”}$

3.2.6. Traffic Density in the Last Period (1-$PM_{TDLP}$)
It is an integrative 1-$PM$ that produces the following set of NL sentences:

- $a_{TDLP_1} \rightarrow \text{“The average road speed in the last period was very low”}$
- $a_{TDLP_2} \rightarrow \text{“The average road speed in the last period was low”}$
- $a_{TDLP_3} \rightarrow \text{“The average road speed in the last period was medium”}$
- $a_{TDLP_4} \rightarrow \text{“The average road speed in the last period was high”}$
- $a_{TDLP_5} \rightarrow \text{“The average road speed in the last period was very high”}$

The validity degrees are obtained by means of the aggregation function $g_{TDLP}$, which calculates the average value of $td$ during a period (76) defined empirically.

3.2.7. Vehicle Speed (1-$PM_{VS}$)
This assertive 1-$PM$ produces, for each detected vehicle, the following set of NL sentences:

- $a_{VS_1} \rightarrow \text{“The vehicle speed is very low”}$
- $a_{VS_2} \rightarrow \text{“The vehicle speed is low”}$
- $a_{VS_3} \rightarrow \text{“The vehicle speed is medium”}$
- $a_{VS_4} \rightarrow \text{“The vehicle speed is high”}$
- $a_{VS_5} \rightarrow \text{“The vehicle speed is very high”}$

3.2.8. Unsafe Speed Conditions (2-$PM_{USC}$)
This integrative 2-$PM$ aggregates the information provided by 1-$CP_{VS}$ during a period of time. Its output includes the following set of NL sentences where we combine crisp quantifying expressions with imprecise quantifying expressions:

- $a_{USC_0} \rightarrow \text{“Zero vehicles speeding”}$
- $a_{USC_1} \rightarrow \text{“One vehicle speeding”}$
- $a_{USC_2} \rightarrow \text{“Two vehicles speeding”}$
- $a_{USC_3} \rightarrow \text{“Three vehicles speeding”}$
- $a_{USC_4} \rightarrow \text{“Four vehicles speeding”}$
\( a_{USC} \rightarrow \text{“Several vehicles speeding”} \)
\( a_{USC} \rightarrow \text{“Many vehicles speeding”} \)

The validity degrees are obtained by means of the aggregation function \( g_{USC} \), which is based on the \( \alpha \)-cuts method proposed by [10]. For example, using the validity degree \( w_{VS} \) of “The vehicle speed is very high”, we calculate the percentage of vehicles with a very high speed contained at each \( \alpha \)-level \( (N_\alpha) \) by means of Eq. 1, with \( \alpha \in A = \{0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9\} \).

\[
N_\alpha = \frac{1}{n} \sum_{i=1}^{n} F_\alpha(w_{VS})
\]

where:

\[
F_\alpha(w_{VS}) = \begin{cases} 
1 & \text{if } w_{VS} > \alpha \\
0 & \text{if } w_{VS} \leq \alpha 
\end{cases}
\]

Then, we calculate the membership degree of each \( N_\alpha \) to each element of the set of linguistic quantifiers: \( \{Q_0, \ldots, Q_6\} = \{\text{Zero, One, Two, Three, Four, Various, Many}\} \), e.g., \( \mu_{Q_3}(N_\alpha) = \text{Three}(N_\alpha) \). Fig. 4 shows these linguistic labels defined on the domain of the number of vehicles being \( n \) a empirically calculated maximum.

3.2.9. Level Of Service (2-PM\text{LOS})

The level of service (LOS) is a measure used by traffic engineers to determine the effectiveness of elements of transportation infrastructure. LOS is most commonly used to analyze highways by categorizing traffic flow with corresponding safe driving conditions. The Highway Capacity Manual [17] distinguishes between six levels of service: A, B, C, D, E, and F. Therefore, we have defined an assertive output \( CP (y_{LOS}) \) that identifies these six levels having the following set of possible sentences:

\( a_{LOS_1} \rightarrow \text{“The level of service is A. Free-flow operation”} \)
Los2 → “The level of service is B. Reasonably free flow; the ability to maneuver is only slightly restricted and the effects of minor incidents still are easily absorbed”

Los3 → “The level of service is C. Stable flow, speeds at or near free-flow and queues may form”

Los4 → “The level of service is D. Approaching unstable flow, speeds decline slightly with increasing flows while density increases more quickly”

Los5 → “The level of service is E. Unstable flow, with operation near or at capacity and no usable gaps in the traffic stream”

Los6 → “The level of service is F. Forced or breakdown flow, queues form behind breakdown points and demand is greater than capacity”

This 2-PM has two 1-CPs as inputs: the traffic density (1-CPTD) and its trend (1-CP_TDT).

The aggregation function (gLos) calculates, at each time instant (t), the value of the validity degrees for each sentence based on the previous validity degrees (time instant t – 1) and current input CPs. Therefore, the aggregation function is a FFSM. For a more detailed description of this paradigm and its applications, the interested reader could see our previous papers [4, 3, 1, 24]. Fig. 5 shows how we use a FFSM to define constraints on the possibilities to change of Los. Using this state diagram, we identify 16 fuzzy rules: 6 rules (R_{ii}) to remain in each Los and other 10 rules (R_{ij}) to change between different Los. This rule base is defined using expert knowledge based on the descriptions of the Highway Capacity Manual [17], which links terms related to traffic density and its evolution along time. It is clear how the system evolution is given by the traffic density, which has a different associated linguistic term for each Los; and its trend, which governs the change to a better or worse state if the density trend is negative or positive, respectively. These rules are listed as follows:

\[ R_{11}: \text{IF } a_{Los}, \text{ AND } a_{TD}, \text{ AND } a_{TDT} \text{ THEN } a_{Los} \]
\[ R_{22}: \text{IF } a_{Los}, \text{ AND } a_{TD}, \text{ AND } a_{TDT} \text{ THEN } a_{Los} \]
\[ R_{33}: \text{IF } a_{Los}, \text{ AND } a_{TD}, \text{ AND } a_{TDT} \text{ THEN } a_{Los} \]
\[ R_{44}: \text{IF } a_{Los}, \text{ AND } a_{TD}, \text{ AND } a_{TDT} \text{ THEN } a_{Los} \]
\[ R_{55}: \text{IF } a_{Los}, \text{ AND } a_{TD}, \text{ AND } a_{TDT} \text{ THEN } a_{Los} \]
\[ R_{66}: \text{IF } a_{Los}, \text{ AND } a_{TD}, \text{ AND } a_{TDT} \text{ THEN } a_{Los} \]
\[ R_{12}: \text{IF } a_{Los}, \text{ AND } a_{TD}, \text{ AND } a_{TDT} \text{ THEN } a_{Los} \]
\[ R_{23}: \text{IF } a_{Los}, \text{ AND } a_{TD}, \text{ AND } a_{TDT} \text{ THEN } a_{Los} \]
R34: IF aLOS_3 AND aTD_3 AND aTDT_3 THEN aLOS_4
R35: IF aLOS_4 AND aTD_4 AND aTDT_3 THEN aLOS_5
R36: IF aLOS_5 AND aTD_5 AND aTDT_3 THEN aLOS_6
R37: IF aLOS_6 AND aTD_5 AND aTDT_3 THEN aLOS_1
R38: IF aLOS_1 AND aTD_6 AND aTDT_3 THEN aLOS_2
R39: IF aLOS_2 AND aTD_6 AND aTDT_3 THEN aLOS_3
R40: IF aLOS_3 AND aTD_6 AND aTDT_3 THEN aLOS_4
R41: IF aLOS_4 AND aTD_6 AND aTDT_3 THEN aLOS_5
R42: IF aLOS_5 AND aTD_6 AND aTDT_3 THEN aLOS_6
R43: IF aLOS_6 AND aTD_6 AND aTDT_3 Then aLOS_1

Where:

- The first term in the antecedent computes the previous validity degree of the sentence aLOS_i, i.e., wLOS_i. With this mechanism, we only allow the FFSM to change from the LOS i to the LOS j (or to remain in LOS i, when i = j). For example, in R11, it is computed the validity degree of the sentence “The level of service is A. Free-flow operation” (wLOS_1).

- The second term in the antecedent describes the constraints imposed on the Traffic Density 1-CP_TD, e.g., “the traffic density is extremely low” (wTD_1).

- The third term in the antecedent describes the constraints imposed on the Traffic Density Trend 1-CP_TDT, e.g., “the traffic density is keeping constant” (wTDT_1).

- Finally, the consequent of the rule defines the next LOS. To calculate the validity degrees of the sentences associated with each LOS j (wLOS_j), a weighted average using the firing degree of each rule R_{ij} (\phi_{ij}) is computed as defined in Eq. 5:

\[
 w_{LOS_j} = \frac{\sum_{i=1}^{6} \phi_{ij}}{\sum_{i=1}^{6} \sum_{j=1}^{6} \phi_{ij}}
\]  

(5)

where \phi_{ij} is calculated using the minimum for the AND operator.

Note that each rule of this set, is therefore, a complete linguistic expression as can be seen in the following expanded expression of the rule R_{11} to remain in the LOS A: “If the previous level of service was A, and the traffic density is extremely low, and the traffic density is decreasing. Then, the current level of service is A. Free-flow operation”.

3.2.10. Level Of Service Trend (2-PM_{LOST})

During the design of this derivative 2-PM, we explored different ways of expressing linguistically the perception of change and, therefore, how to calculate their validity degree. This derivative 2-PM has the Level Of Service (2-CP_{LOS}) as input. The output CP y_{LOST} includes four types of NL propositions for each LOS i at each time instant t:

- a_{LOST_{1i}} \rightarrow “The level of service is keeping constant in level i”
- a_{LOST_{2i}} \rightarrow “The level of service is changing from level i to level j”
- a_{LOST_{3i}} \rightarrow “The level of service of the road has changed to level j”
- a_{LOST_{4i}} \rightarrow “The level of service of the road has returned to level i. The change has not been completed”

Here, the aggregation function g_{LOST} calculates the trend of each LOS by analyzing the derivative of the validity degrees of each sentence a_{LOS_i} (wLOS_i). At each time instant we determine if a certain LOS i is decreasing (D_{LOST_i}), keeping constant (KC_{LOST_i}), or increasing (I_{LOST_i}) by fuzzifying its derivative.
We also defined that there has been a change in the level of service (denoted by $C_{ij}$), when a certain level $i$, which had a higher validity degree than other level $j$, becomes smaller than $j$. This binary indicator takes value 0 when there is not a change and 1 when a change is produced. After careful experimentation, we have defined the validity degrees of each type of sentence at each time instant $t$ as follows:

$$w_{\text{LOST}_i}[t] = \min (KC_{\text{LOST}_i} \cdot 1 - C_{ij}).$$

A LOS is keeping constant when the previous level was the same ($C_{ij} = 0$) and it is keeping constant ($KC_{\text{LOST}_i}$).

$$w_{\text{LOST}_{ij}}[t] = \min (D_{\text{LOST}_i} \cdot I_{\text{LOST}_i} \cdot 1 - C_{ij}).$$

A level $i$ is changing to a level $j$ when $i$ is decreasing ($D_{\text{LOST}_i}$), $j$ is increasing ($I_{\text{LOST}_i}$), and the change between levels has not been completed yet ($C_{ij} = 0$).

$$w_{\text{LOST}_{ij}}[t] = C_{ij}.$$ A LOS has recently changed to a level $j$ when the LOS of the previous time instant was different to the actual one ($C_{ij} = 1$).

$$w_{\text{LOST}_i}[t] = \min (I_{\text{LOST}_i}, D_{\text{LOST}_i}, w_{\text{LOST}_{ij}}[t - 1]).$$

In some many cases, a LOS could have been changing but the change was not completed. For example, the LOS could be changing from level $E$ to level $F$ but suddenly the density decreases and the change is stopped, keeping at level $E$. This case is recognized when the system was changing from level $i$ to level $j$ ($w_{\text{LOST}_{ij}}[t - 1]$) but the level $i$ starts to increase ($I_{\text{LOST}_i}$) and the expected level $j$ starts to decrease ($D_{\text{LOST}_j}$).

3.2.11. Level Of Service in the Last Period (2-PM$\text{LoSLP}$)

Here, we experiment with another example of 2-PM. This integrative 2-PM has the Level Of Service (2-CP$\text{LoS}$) as input $CP$. The output CP $y_{\text{LoSLP}}$ includes four types of NL propositions for each LOS $i$ that summarize the amount of times that each LOS has been activated during a certain period of time:

$$a_{\text{LoSLP}_i} \rightarrow \text{"In the last period, the level of service has never been } i\text{"}$$
$$a_{\text{LoSLP}_i} \rightarrow \text{"In the last period, the level of service has been few times } i\text{"}$$
$$a_{\text{LoSLP}_i} \rightarrow \text{"In the last period, the level of service has been sometimes } i\text{"}$$
$$a_{\text{LoSLP}_i} \rightarrow \text{"In the last period, the level of service has been many times } i\text{"}$$

These sentences are based on the summarizers proposed by [27]: “Q the LOS has been $i$”. Where $Q$ is a fuzzy quantifier [28] applied on the cardinality of the perception “the LOS has been $R$”. And $R$ is the summarizer, in this case the set of possible LOS. The set of linguistic labels for each quantifier are uniformly distributed trapezoidal SFPs denoted by the expressions never, few times, sometimes and many times.

This information is really important to summarize traffic behavior in a certain amount of time because it allows to compare the state and trend of traffic in a specific road whose study is interesting to obtain conclusions such as checking the need to redirect the traffic. The aggregation function ($y_{\text{LoSLP}}$) calculates the validity degrees for each sentence based on the cardinality values ($Card$) of the validity degrees of each sentence $a_{\text{LoS}}$, ($w_{\text{LoS}}$) during the desired period duration:

$$a_{\text{LoSLP}_i} = \text{never} [Card(w_{\text{LoS}})]$$
$$a_{\text{LoSLP}_i} = \text{few times} [Card(w_{\text{LoS}})]$$
$$a_{\text{LoSLP}_i} = \text{sometimes} [Card(w_{\text{LoS}})]$$
$$a_{\text{LoSLP}_i} = \text{many times} [Card(w_{\text{LoS}})]$$

3.3. Expression Module

Apart from the goal of obtaining suitable texts to be showed to drivers, the linguistic reports can be used by traffic experts with the aim of understanding changes in traffic and foreseeing its future behavior. Using the set of available $CP$s in the GLMP, namely, the evolution of the LOS, vehicles speed, road speed trend, extraordinary speed conditions and so on, the developed application provides two different types of linguistic description reports: an specific report which describes the instantaneous state of the traffic, and a periodical report that summarizes traffic behavior throughout a specific period of time. In both cases, we have applied basic report templates, see in [2] an example of a template that change the structure of the report depending on the validity degrees of the sentences.
3.3.1. Specific report

A specific time instant or eventual report about the traffic behavior is given. The periodicity of these reports depends on the final user’s needs (one minute, five minutes, ten minutes, etc.). Each report informs about the LOS trend (changing, recently changed or keeping constant), the traffic density and road speed in the last period of time, traffic density trend and road speed trend.

The report template is represented in Fig. 6. It uses the sentences provided by the traffic density and its trend CPs (1-CP_{TD} and 1-CP_{TDT}), the LOS and its trend CPs (2-CP_{LOS} and 2-CP_{LOST}), and the road speed and its trend CPs (1-CP_{RS} and 1-CP_{RST}). The sentences with the highest validity degree are chosen at each time instant for each CP. This type of information shows us that it is possible that some CPs are changing while the LOS is keeping constant, e.g., the traffic density can be increasing inside the level C but this does not mean that LOS is changing from level C to level D. This is the difference between the changes of the traffic density during a certain LOS and the changes between different LOS. One possible specific traffic report may be as follows: Currently, the Traffic Density is low and it is increasing. The Level Of Service is changing from level B to level C, stable flow, speeds at or near free-flow and queues may form. The Road Speed is medium and it is decreasing.”
3.3.2. Periodical report

This report summarizes traffic behavior throughout a period of time, e.g., a full day or other sets of periods that can give relevant information about the traffic progress. Traffic experts decide how many periods they want to analyze separately, in order to verify the differences existing among them. This type of information allows them to extract conclusions and to implement appropriate measures to improve the quality of traffic (improving infrastructure, notice drivers and so on). For example, one type of differentiation could be analyzing separately sunrise, morning, midday, afternoon, evening and night. In the same way, the final user could decide to segment the day into smaller periods and to extract information about periods of different sizes. The traffic summary report also gives information relative to the average traffic density and the average level of service in each period of time.

Therefore, a different report template must be used. It is represented in Fig. 7 and it uses the sentences provided by the traffic density in the last period $CP_1$ ($1-CP_{TDLP}$), the LOS in the last period $CP_2$ ($2-CP_{LOSLP}$), the road speed in the last period $CP_3$ ($1-CP_{RSLP}$), and the unsafe speed conditions $CP_4$ ($1-CP_{USC}$). Similarly to the specific report, the sentences with the highest validity degree are chosen for each $CP$. One possible global traffic report throughout the afternoon may be as follows: “In the afternoon, the Traffic Density has been medium. The Level Of Service has never been A and B; and sometimes C, D, E and F. The Road Speed has been low. There were not vehicles speeding.”.

4. Experimentation

4.1. Simulated traffic data

In order to deal with a broad number of situation types, we have designed a simulator after analyzing several databases of traffic control centers of important cities, such as Madrid, Valencia, and Sevilla [16]. The simulator is based on the Monte Carlo method where simulated data (number of cars, its size and its speed) follow a normal distribution. Each normal distribution is defined by its mean and its standard deviation, e.g., these parameters vary depending on the period of the day. This simulator allow us to generate data that recreates the traffic behavior in different situation types providing data each five minutes of simulated time. Fig. 8 shows an example of simulation of traffic density and road speed along a typical working day, and includes the validity degrees obtained for the sentences related to the LOS.

In the following, we show several examples of specific (represented with the symbols ■, ▲, •, and ♦ in Fig. 8) and periodic traffic reports associated to these simulated data. Every report, specific or periodic, is accompanied by its reference of time, either the period of the day or the specific time of measurement. Results are consistent and show accurately the simulated situations.

- **Specific reports:**
  
  ■ “At 10:55, the Traffic Density is extremely high and it is decreasing. The Level Of Service is keeping constant in level F, forced or breakdown flow, queues form behind breakdown points and demand is greater than capacity. The Road Speed is low and it is keeping constant”
  
  ▲ “At 16:15, the Traffic Density is medium and it is decreasing. The Level Of Service has changed to level C, stable flow, speeds at or near free-flow and queues may form. The Road Speed is low and it is keeping constant”
  
  • “At 4:10, the Traffic Density is extremely low and it is increasing. The Level Of Service is changing from level A to level B, reasonably free flow, the ability to maneuver is only slightly restricted and the effects of minor incidents still are easily absorbed. The Road Speed is high and it is decreasing”
  
  ♦ “At 4:45, the Traffic Density is extremely low and it is keeping constant. The Level Of Service has returned to level A, free-flow operation. The Road Speed is medium and it is increasing”

- **Periodical reports:**
Figure 8: Graphical representation of the Traffic Density, Road Speed, and the validity degrees of the sentences associated to each Level of Service.

- “In the morning (from 7:00 to 10:00), the Traffic Density has been extremely high. The Level Of Service has never been A and B; few times D; sometimes C and E; and many times F. The Road Speed has been low. There were 4 vehicles speeding”

- “In the afternoon (from 13:00 to 22:00), the Traffic Density has been low. The Level Of Service has never been A and B; and sometimes C, D, E and F. The Road Speed has been low. There were not vehicles speeding”

- “At night (from 22:00 to 7:00), the Traffic Density has been extremely low. The Level Of Service has never been E and F; few times C and D; sometimes B; and many times A. The Road Speed has been medium. There were many vehicles speeding”

- “During the whole day, the Traffic Density has been medium. The Level Of Service has been few times B, D, E and F; and sometimes A and C. The Road Speed has been low. There were many vehicles speeding”

4.2. Video camera real data

In order to check the performance and effectiveness of our application in a real situation, we have used digital image processing techniques to acquire the input data from a video camera. In a previous work
we have already used the images of a video stream, where we used an overhead camera to acquire real images and generate successfully linguistic descriptions about traffic conditions in a roundabout. Here, we used images obtained from the right side of the road. Here, we acquired information from live video recordings, recognizing vehicles and their different features (speed, position and size).

In order to recognize the image content, we considered enter and exit regions, rail regions, image perspective and obstacles that could appear in images. In our model, each vehicle moving along the analyzed road region is characterized by a unique identification number that is assigned after it is first detected passing by any enter region of the road. Fig. 9 shows graphically a typical situation in which we can observe a set of vehicles identified by a number and highlighted with its bounding box.

The specific traffic reports obtained for these two frames were as follows: “Currently, the Traffic Density is high and it is decreasing. The Level Of Service has returned to level D, approaching unstable flow, speeds decline slightly with increasing flows while density increases more quickly. The Road Speed is low and it is decreasing” and “Currently, the Traffic Density is low and it is decreasing. The Level Of Service is keeping constant in level C, stable flow, speeds at or near free-flow and queues may form. The Road Speed is low and it is increasing”.

5. Concluding remarks

During the last few years, we have developed an extension of CTP that allows generating linguistic descriptions of complex phenomena. In this work, our goal consists of exploring a practical application in the field of ITS that led us to design new types of PMs, i.e., we aimed to improve the number of available types of linguistic expressions and therefore the versatility of our technology. Together with the practical application, in this paper, we have contributed to the development a new general approach to produce linguistic descriptions of complex phenomena evolving in time. Specifically, we have designed several 2-PMs demonstrating how to model the meaning of several different linguistic expressions belonging to the specific application domain of language. Moreover, we have showed that depending on the user requirements, our approach allows us to generate a great variety of customizable linguistic reports.

However, there is still too much work to do. Indeed, NL has a limitless potential of meaning. From the theoretical point of view, we will continue exploring how to model the meaning of different linguistic expressions, i.e., new CPs and PMs. From the practical point of view, we will continue developing practical applications, e.g., we will try to set up a complete system for monitoring and control the traffic in a real world scenario. The experiments performed using simulated data have allowed us to demonstrate the great expressiveness of the presented resources. The experiments performed using data obtained from real video images have allowed us to demonstrate the viability of our approach to create real industrial applications.
Acknowledgments
This work has been funded by the Spanish Government (MICINN) under project TIN2011-29827-C02-01.

References


