Scientific Journals
of the Maritime University of Szczecin

Zeszyty Naukowe
Akademii Morskiej w Szczecinie

2016, 47 (119), 176–180
ISSN 1733-8670 (Printed)
ISSN 2392-0378 (Online)
DOI: 10.17402/166

Received: *29.04.2016*
Accepted: *19.07.2016*
Published: *20.09.2016*

Methods for building a knowledge base for automatic communication in shipping

Anna Wójcik✉, **Zbigniew Pietrzykowski**

Maritime University of Szczecin
1–2 Wały Chrobrego St., 70-500 Szczecin, Poland
e-mails: {a.wojcik; z.pietrzykowski}@am.szczcin.pl
✉ corresponding author

Key words: e-navigation, automation of communication processes, knowledge base, inference rules, classical predicate calculus, computing with words

Abstract

One of the most significant phases for automation of communication processes in shipping is building a knowledge base for inference processes. Communication processes include: exchange of information, perception of communication and interaction between navigators. Computing with words has been used to represent inference processes covering imprecise concepts that are characteristic of natural languages. Elements of classical predicate calculus were adopted as a basic form of writing inference rules. Methods for constructing a knowledge base were chosen. The knowledge base architecture was proposed. This article also presents examples of inference rules in a knowledge base for automatic communication in shipping.

Introduction

The process of ship conduct requires constant information exchange and processing of navigational information. Human error is one of the main causes of shipping accidents. They can result from a lack of the information necessary to make decisions, or their misinterpretation, but also from an overload of information, hindering the processing and selection of information relevant to the decision-making process. Navigators, to assess the situation and make decisions, use shipboard devices and systems, and means of verbal communication. Verbal communication comprises of the exchange of information, including the acquisition of additional information and, if appropriate, negotiations.

Problems with correct verbal communication are mostly associated with the failure to communicate, misunderstanding messages, a wrong choice of message or misinterpretation of the information exchanged. These errors can be due to navigators' mental and physical state, particularly with regard to stress and fatigue.

One of the proposed solutions is the automation of communication processes in shipping, incorporating such aspects as selective acquisition of information, including other navigators' intentions and their interpretation, and taking into account interactions between the navigators, their negotiations in particular. The automation of communication processes requires a knowledge base to be developed for inference processes to take place. Such a knowledge base has to allow for the complexity of communication processes, including negotiations, and the specifics of marine navigation.

The construction of the knowledge base

When creating knowledge bases and generating rules, it is necessary to define the properties these bases and rules should have. A knowledge base may represent a particular scope of knowledge in a variety of ways. For automatic communication in shipping we have chosen a rule-based knowledge base, i.e. the rule base. The rules have the form "If ... then ..." and take into account assumptions of fuzzy logic

to include imprecise terms of natural language. For the building of knowledge bases and rule bases, the following properties are generally pointed out (Piegat, 1999):
- dependence of the number of rules on the number of inputs and fuzzy sets;
- completeness;
- compatibility (consistency);
- continuity;
- redundancy.

The rule base in question does not have to meet all of these requirements.

In the case of a knowledge base containing rules of fuzzy logic, one essential feature is the exponential dependence of the number of rules on the number of inputs and the number of fuzzy sets in the model. This entails a rapid increase of model complexity that, on the one hand, can raise the accuracy of the mapping of the real system, but on the other hand, can require more information to determine the parameters of the membership function. The base under construction quickly becomes complex, which can lead to difficulties in maintaining its other vital features.

The rule base is defined as complete if the base can assign a certain state of the output to every state of the input. The database is not complete if there are such input states to which this base cannot assign any output state. A model with a complete database of rules more accurately represents the operation of a real system, but in the early stages of rule generation and testing the base is incomplete. Due to the dynamics of verbal communication problems and the diversity of navigational situations, the knowledge base under construction is referred to as an incomplete base, which in specific cases has a capability of generating new rules that are then added to the base.

The rule base is consistent if it does not contain conflicting rules, i.e. rules based on the same preconditions but with different conclusions. This may be due to rule generation errors or the ambiguity of the real system itself.

The rule base in the fuzzy inference system is continuous if it has no neighbouring rules: R_jR_k with respective conclusions, h_j, h_k, whose product is empty, i.e. $h_j \cap h_k = \varnothing$. The continuity of the rule base is recommended but not necessary. Discontinuity means rapid changes in the values of conclusions.

The redundancy of a database occurs when it contains two or more identical rules that were created, for example, at various stages of the base construction. The knowledge base should not have redundant rules.

Stages of knowledge base construction

Building a knowledge base takes place in the steps shown in Figure 1 (Kent, 2000)

Identification involves determining the characteristics of the problem being solved. The problem itself and the scope it relates to are defined precisely.

Representation means working out a manner of knowledge representation. Once the problem is analysed and understood, we can determine the information and data needed to solve the problem. The gathering of necessary data starts at this stage.

Formalisation is the stage in which structures that organise the knowledge are designed, that is key concepts, rules and relationships are translated into formal language. The syntax and semantics of the language are developed, then all the basic concepts and relations necessary to solve the given problem are established.

Implementation means the formulation of the rules or framework comprising the knowledge. At this stage, the formalised knowledge of the previous stage is combined and reorganised so as to make it compatible with the characteristics of the information flow of a given problem. The resulting set of rules or frames and the associated control structure create a prototype program.

Testing is the last stage of building a knowledge base. It involves checking the system rules or frames. The rules and relations are tested for generation of correct responses.

This article will cover the elements of the first three stages.

The processes of inference

Automatic communication in shipping encompasses three basic processes:
- identification of a navigational situation;
- classification of the navigational situation (navigational situation recognition);
- communication processes.

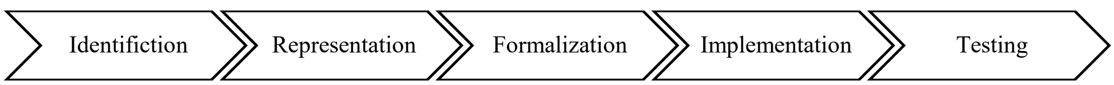

Figure 1. Stages of knowledge base construction

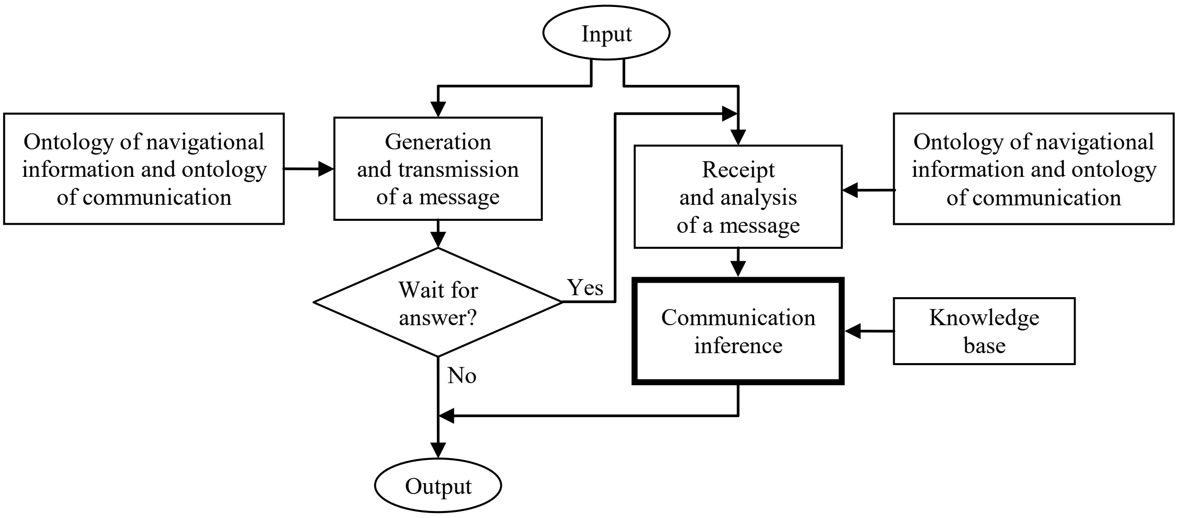

Figure 2. Schematic communication process (Pietrzykowski et al., 2015)

The first process aims at identifying potential risks arising from an encounter. Based on the analysis of the current situation, including parameters CPA and TCPA, it is identified either as safe or potentially dangerous. This step of the system operation is referred to as the preliminary inference.

The classification of a navigational situation comes next, based on the Collision Regulations. Decisions are made concerning which ship has the right of way, the need for manoeuvres to be performed by the navigators conducting their vessels, and the possible need for establishing verbal communication. The process of inference at this stage is inference proper.

Finally, communication processes including reception and analysis of information, and inference on that basis, as well as generating and sending feedback to the target vessel's navigator or automatic communication system, are shown in Figure 2.

Let us consider two cases:

- the system identifies a situation that requires establishing communication;
- the system is called by another ship (by receiving a message).

After the message is interpreted using the ontologies of navigational information and communication, the data is transmitted to the Communication Inference block. The input data to this block are messages received from another ship or a coast station in text form (natural language) as well as data from shipboard systems. If the input data is crisp, the performed inference is of classical type, using the knowledge base. If the input data includes imprecise terms, the system makes use of inference using fuzzy logic and methods of computing with words (Figure 3).

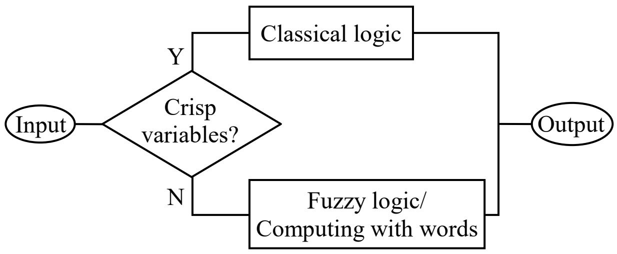

Figure 3. Diagram of inference processes in communication (Wójcik, Hatłas & Pietrzykowski, 2016)

The system works on the principles of fuzzy inference systems with a properly constructed knowledge base. Fuzzy systems can be divided into four basic elements, shown in Figure 4.

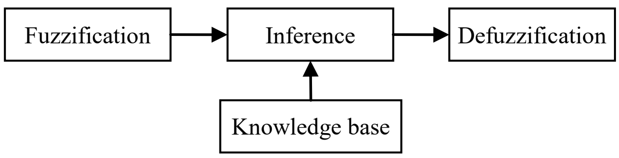

Figure 4. Diagram of fuzzy inference processes

Input values of linguistic variables are expressed by fuzzy sets in the fuzzification block. The input values are fuzzified by assigning a membership function assuming a value from the interval [0, 1] to a variable. In the event of a crisp value occurring at the input, the fuzzification of a singleton type, the membership function takes only the values 0 or 1.

In the inference block, those rules from the knowledge base are executed whose preconditions are met by the data of the previous block. The system calculates a fuzzy set resulting from the operation of the model used, e.g. Mamdani or Takagi–Sugenau model.

The defuzzification block is responsible for mapping a fuzzy set that came out of the inference block onto a value that will be an output of the fuzzy system.

The knowledge base is a collection of data relating to a specific area with the logic rules allowing for inference.

The knowledge base created for automatic communication in shipping is expected to allow inference on the basis of crisp and fuzzy input data, with possible use of computing with words. To this end, we propose to build a relational knowledge base. The rules in this base principally take the following form (Pietrzykowski et al., 2011):

rule: <preconditions> → <conclusions>

where: <preconditions> (often called premises) indicate when a given rule may be applied, and <conclusions> denote the effect of rule application, which may be a logical value, a decision or action.

The rules often contain a few preconditions, linked by logical connectives, leading to a single conclusion, and are written as follows:

r: IF p_1 AND p_2 AND p_3 … AND p_n THEN h

where: r means a rule, $p_1, p_2, p_3, \ldots, p_n$ – preconditions, and h denotes a conclusion.

Preconditions may be built using propositional logic, attributive logic or classical predicate calculus, i.e. first-order logic.

Methods of building the knowledge base

When using propositional logic, it is assumed that each sentence can be assigned one of two logical values: true or false. This causes limitations in knowledge representation. The syntax features propositional variables (sentences), sentence-forming functors (i.e. propositional connectives) and auxiliary symbols (e.g. parentheses). Inference systems may, for example, take the form of decision tables, decision lists, decision rules with control rules or a decision tree (Ligęza, 2006).

In case of attribute logic using attributes and values, the attribute is used to write down some properties of objects and the system to which they are applied. A set of characteristic attributes is chosen, with certain values assigned to them. This method is most often used to define facts concerning a system, specifications and properties of programs and their components.

The classical predicate calculus is the most popular way of recording the statements used to express the knowledge described in natural language. It allows the use of variables, conditions and quantifiers, which make it possible to formalize complex knowledge. This method was chosen as the basis for writing down the rules in the system of automatic communication and will be described in detail in the next section.

Depending on the selected method of statement notation, the activation of rules is dependent on the fulfilment of conditions and the choice of a rule by the inference machine. The rules are selected and checked by the inference machine using different predefined algorithms. These may be methods of checking the rules for selected properties, in a serial or parallel manner. Individual rules are then checked for the fulfilment of preconditions.

The classical predicate calculus

For this proposed system of automatic communication in shipping, classical predicate calculus has been adopted as the primary method of rule notation. This choice is dictated by the nature of ambiguous input, which may take the form of imprecise terms of natural language.

The term predicate means a property or relationship of certain objects which are its arguments. N is the argument predicate over a class of individuals; X denotes the mapping

$$P: X^n \rightarrow \{F, T\}$$

that to every n-element string of individuals assigns a Boolean value, false – F or true – T. The basic statement in the predicate calculus assumes the form $P(x_1, x_2, x_3, ..., x_n)$ and means that the relation P is maintained for objects $x_1, x_2, x_3, ..., x_n$. The individual variable x is bound in an expression if x is a variable of the quantifier $\forall x$ (for every x), or $\exists x$ (there exists x), otherwise the variable x in this expression is free. The expression containing no free variables is called a closed form.

A propositional expression is interpreted as a closed form. If a predicate expression contains a free variable, then it is interpreted as a relation in a certain class of individuals. The expression of predicate calculus is defined as true in a given interpretation if any substitution of class individuals in place of variables leads to a true sentence. Formulas of classical predicate calculus are constructed in the same manner as in propositional calculus, but with the addition of variables and quantifiers.

The knowledge base under construction will contain inference rules and metarules (rules instructing how to use other rules). Let us consider an encounter of two vessels A – Alpha and B – Beta, proceeding on reciprocal courses (Pietrzykowski et al., 2011), for which we will write down an example form of an inference rule (Figure 5).

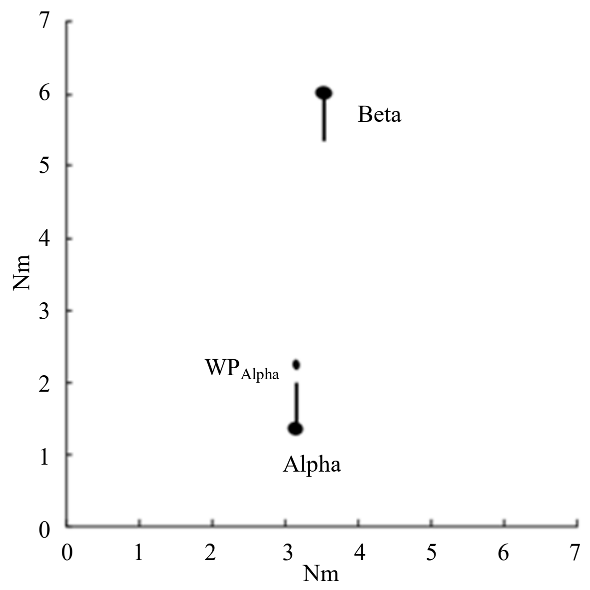

Figure 5. An example navigational situation

According to the regulations in force, the ships should pass each other port-to-port. For the presented situation, part of verbal communication of the ships' navigators may go like this:

A: Our CPA is 0.2 Nm.

A: I intend to alter my course to starboard soon and cross ahead of you at a safe distance.

B: Ok, cross ahead of me at a distance farther than 1.2 Nm.

Inference processes are considered from the viewpoint of the system on the ship Alpha.

An example notation of the rule for the above dialog has this form:

$\forall$A, B $\in$ ships
[relative bearing (A, 5) $\wedge$ aspect (B, 5)
$\wedge$ CPA (A, B) $\wedge$ wp (A)]
$\rightarrow$ [manoeuvre (A, passing)
$\wedge$ manoeuvre (B, standing on)
$\wedge$ inf o preparation]

where:
relative bearing (A, 5) – angle A $\leq 5°$;
aspect (B, 5) – aspect B $\leq 5°$;
CPA(A, B) – CPA(A, B) $\leq CPA_{Limit}$;
wp(A) – information on next way point.

This notation represents a formal record of rules of inference and metarules that will be generated in further research. The notation takes into account the accepted methods of knowledge representation in the inference system for automatic communication.

Conclusions

The verbal communication discussed herein involves the exchange of information, including acquisition of additional information and, where appropriate, negotiations. Problems of normal vocal communication are most often associated with failure to establish radio communication, misunderstanding of messages, improper choice of message (wording) or incorrect interpretation of information being exchanged.

The proposed system of automatic communication can help reduce these errors. The system is built on a rule base, a specific type of knowledge base. The rules are written in line with classical predicate calculus using fuzzy logic for imprecise terms of natural language.

The article characterises stages of building a knowledge base and identifies the properties of the rule base required for inference processes. Selected elements of the knowledge base are presented along with an example of formal notation of inference rules. This kind of notation allows us to develop inference rules and the governing rules, often referred to as metarules. Currently, the test environment is being prepared to verify the created rules.

Acknowledgments

This research outcome has been achieved under the research project No. 14/MN/ITM/16 financed from a subsidy of the Ministry of Science and Higher Education for statutory activities of Maritime University of Szczecin.

References

1. Kent, A. (2000) *Encyclopaedia of Library and Information Science: Volume 68 (Supplement 31)*. CRC Press.
2. Piegat, A. (1999) *Modelowanie i sterowanie rozmyte*. EXIT.
3. Pietrzykowski, Z., Hatłas, P., Wójcik, A. & Wołejsza, P. (2015) *Subontology of communication in the automation of negotiating processes in maritime navigation*. Conference proceedings MTE 2015.
4. Pietrzykowski, Z., Hołowiński, G., Magaj, J. & Chomski, J. (2011) Automation of Message Interchange Processes in Maritime Transport. *International Journal on Marine Navigation and Safety of Sea Transportation* 5, 2.
5. Wójcik, A., Hatłas, P. & Pietrzykowski, Z. (2016) Modelling Communication Processes in Maritime Transport Using Computing with Words, TST2016.
6. Ligęza, A. (2006) *Logical Foundations for Rule-Based Systems*. Studies in Computational Intelligence, Volume 11.