Case Learning in CBR-Based Agent Systems for Ship Collision Avoidance

Liu, Yuhong; Yang, Chunsheng; Yang, Yubin; Lin, Fuhua; Du, Xuanmin

Publisher's version / Version de l'éditeur:

The Proceedings of the 12th International Conference on Principles of Practice in Multi-Agent Systems 2009, 2009-12-16

Case Learning in CBR-Based Agent Systems for Ship Collision Avoidance

Yuhong Liu[1], Chunsheng Yang[2], Yubin Yang[3], Fuhua Lin[4], and Xuanmin Du[5]

[1] Merchant Marine College of Shanghai Maritime University, Shanghai 200135, China
[2] Institute for Information Technology, National Research Council Canada
[3] State Key Laboratory for Novel Software Technology, Nanjing University, China
[4]School of Computing and Information Systems, Athabasca University, Canada,
[5] Shanghai Marine Electronic Equipment Research Institute, Shanghai 201108, China

Abstract. With the rapid development of case-based reasoning (CBR) techniques, CBR has been widely applied to real-world applications such as agent-based systems for ship collision avoidance. A successful CBR-based system relies on a high-quality case base. Automated case creation technique is highly demanded. In this paper, we propose an automated case learning method for CBR-based agent systems. Building on techniques from CBR and natural language processing, we developed a method for learning cases from maritime affair records. After reviewing the developed agent-based systems for ship collision avoidance, we present the proposed framework and the experiments conducted in case generation. The experimental results show the usefulness and applicability of case learning approach for generating cases from the historic maritime affair records.

Keywords: case-based reasoning, multi-agent system, ship collision avoidance, maritime affair records, case learning, case base management/updating.

1. Introduction

Human error is one of the most important factors in maritime accidents. In particular, it was a root of ship collision avoidances. To improve the navigation safety and avoid human error, an amount of research work [1-6] has focused on developing intelligent systems for collision avoidance. Yang et al. [4] developed a rule-based expert system based on the navigators' experiences and applied it to an integrated navigation system as a decision-making support system for collision avoidance. Similarly, the authors in [2] [3] developed an intelligent decision support system for ship collision avoidance. These intelligent systems were developed based on rule-based reasoning techniques. The rules were created or obtained from traffic regulations, or encounter scenarios, or navigation theories. Therefore, such rules cannot fully mimic the human's ship-handling behavior and experience, which is the most important factor in ship-handling for collision avoidance. This is why it is difficult to apply these research results to practical navigation systems.

To overcome the shortcomings in rule-based reasoning systems, we have started to look into applying Case-Based Reasoning (CBR) and agent techniques to ship collision avoidance [15][17][18]. CBR is one of the reasoning paradigms and is a feasible and efficient way to the problems which are difficult to be solved with traditional methods such as model-based reasoning. CBR-based approach has been widely applied to different real-world applications such as diagnostic, design, and decision-making support [7-12]. Moreover, we are also looking into applying agent techniques to ship collision avoidance as well. A ship, navigating in an open and dynamic environment, can be looked as a rational and intelligent agent. Ships with navigators can detect the changes of the environment, collect the information of other ships, judge the dangerous degree of current situation, make decisions by using some knowledge, and take actions to avoid the collision with other ships or obstacles. To facility this research, we have developed a multi-agent system for ship collision avoidance by using agent and CBR techniques. The agent in this system [17] was implemented with CBR-based decision-making support for collision avoidance [18].

When we develop a CBR-based system for any applications, a significant challenge that faces us is case generation. Without a high-quality case base, it is impossible for a CBR-based system to function well for solving the given problems. It is a challenge because different domain applications require different approaches for case generation. For example, the authors in [16] developed a methodology to automatically generate cases from the historic maintenance database for diagnostic CBR systems. In this study, we looked into the historic maritime affairs records which were collected in navigation over many years. These records reflect either instructive and successful cases or edifying and failing cases. These cases are a valuable resource to generate cases for CBR-based collision systems. To this end, we have collected many famous maritime cases from Asia and Europe from 1976 to 2006. These cases are documented in an unstructured text format and in different language, mainly in Chinese. To efficiently generate cases from these unstructured text records, we developed a framework, focusing on Chinese text format, by using techniques from natural language processing (NLP) and CBR. In this paper, we present the developed techniques and some preliminary results.

The next section gives an overview on the developed CBR-based multi-agent system for ship collision avoidance; following that, we mainly present a framework for generating cases from maritime affair records; Section 4 introduces the experiments along with some preliminary results; the final section is conclusions and future work.

2. Overview of the Developed CBR-Based Agent Systems

In order to conduct the research for collision avoidance, we developed a multi-agent system [17] for simulating real navigation environments. It consists of two types of agents: control agent and function agent. The control agent contains system agent and union agent; and function agent is either ship agent or VTS (Vessel Traffic Service) agent. In general, the control agents manage the function agents, including information maintenance, agent communication, task partition and assignment, resource distribution and administration, conflict reconcilement, etc. The function agent performs CBR-based reasoning for collision avoidance by using the information from control agents and the environment data.

The function agents such as Ship agent and VTS agent are implemented following the BDI (Beliefs, Desires, and Intentions) model [19]. A BDI agent is a particular type of bounded rational software agent, imbued with particular mental attitudes, such as Beliefs, Desires and Intentions. The BDI model has some philosophical basis in the Belief-Desire-Intention theory of human practical reasoning. We applied CBR to BDI model for modeling human reasoning on collision avoidance. With the help of CBR, the architecture of CBR-based Ship agent and VTS agent is designed as shown in Figure 1. Basically, a CBR-based BDI agent implemented consists of two types of components: BDI function components and CBR function components. BDI function components such as communication, action trigger, model base, provide ability to interact with other agents. CBR function components consist of three main components for performing CBR reasoning: problem description, case bases, and case learning. The problem description component creates a description for a collision situation based on real-time navigation data. These data include static information

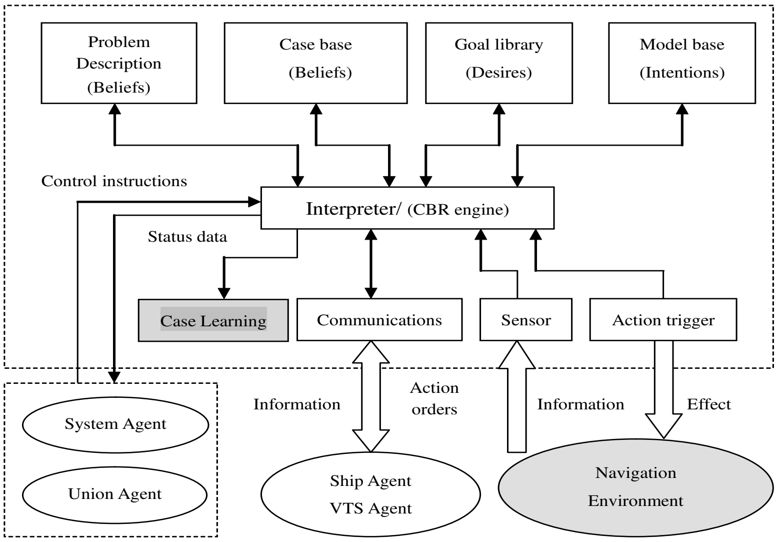

Figure 1: The detail design of the CBR-based agent system

(such as ship type, ship length and sea gauge), dynamic information (such as course, speed and position), and navigation information (such as the relative course and speed, azimuth, distance, DCPA, TCPA, encounter situation, and collision risk). Case base stores cases with a given presentation and an index structure. Once a collision problem is defined from the problem description component, a case retrieval algorithm is used to retrieve similar cases from the case base. The case with maximal similarity is selected as the proposed solution for the current collision problem. The case learning component is a core of CBR-based agent system. The main task is to automatically generate cases from maritime affaire records or a real-time ship-handling simulation. We will present this component in detail in the following Section.

3. Case Learning

To make CBR function well for collision avoidance, a high-quality case base is required. Cases can be generated either in running time or at the initial stage of system development. For collision avoidance it is not feasible to generate cases in running time because of cost and safety. In our previous work [18], we proposed to create cases from ship-handling simulators. By analyzing the ship-handling trajectories we can create some cases for designated encountering cases. Another feasible and effective way to generate cases is using maritime affaire records collected in the navigation history. This section presents a framework to automatically generate cases from historic maritime affair records. We first give some preliminaries for a collision avoidance case based on navigation practice. Then we present the proposed framework for case learning.

3.1 Preliminaries

Ship collision avoidance is a dynamic process having a close connection with the sea, the ship, the human, and the environment, and involving much information and changes during a period. In order to describe our framework, we first give out some definition on ship encountering situation, viewpoint of navigation, maritime affair data, case and case base.

Definition 1, **Encounter Scene (ES)**: a well-defined data structure. It is used to record the environment information (EI), the basic information (BI) of each ship, the relative information (RI) between own ship and each target ship and the proposed actions (PA) at a given time point. That is:

$$ES = < EI,\ BI,\ RI,\ PA > \tag{1}$$

Definition 2, **View Point (VP)**: during the ship collision avoidance, we label one of the encountering ships as the own ship (OS) and the others as target ships (TS). And then we select a time point T and record the encounter scene (ES) at this moment. VP is denoted as:

$$VP = < OS,\ TS,\ T,\ ES > \tag{2}$$

Definition 3, **Case base (CB)*:*** A case (c) is defined as $c = \{p_c, s_c, m_c\}$. P_c denotes a set of problem attributes, which describes a collision problem, and a set of VP; S_c is a set of solution attributes, either a single action or several actions for avoiding a collision; m_c contains all attributes related to case base maintenance, including redundancy, inconsistency, success times, collision times, successful actions, and failed actions. Let CB denote a case base, where $CB \supseteq \{c_1, c_2 \ldots\ldots c_i, \ldots c_n\}$.

Definition 4, **Maritime Affair Database (MD)*:*** For a given maritime affair record noted as ***md****i*, it contains implicitly or explicitly environment description (ED) (sea state, weather condition, and visibility), ship information description (SID) (encounter ships, ship name, type, length, draft, cargos and operation condition), and the collision or collision avoidance procedure description (PD) (the dynamic operation process and ship-ship relative information). For the collected maritime affair data, we denote it as a database, ***MD***, where $MD \supseteq \{md_1, md_2 \ldots\ldots md_i, \ldots md_q\}$.

3.2 The Method of Case Base Creation

From the above definitions, our task is to create a CB from a given MD. The task of automating the case base creation is accomplished by the following three processes:

- Identifying a collision problem and its solutions
- Creating a template case
- Updating a case base

A. Identifying a collision problem and its solutions

The task of this process is to find p_c and the solution s_c for a given md_i in ***MD.*** In this work, the collected ***MD*** is unstructured Chinese text. Such Chinese text format makes the work more complicated. Unlike English or other western languages, Chinese is character based, not word based. There are no "blank spaces" serving as word boundaries in Chinese sentences [20] [21]. In order to obtain ED, SID and PD information from a given md_i, we first conduct Chinese word segmentation, and then perform semantic analysis according to a selected view point (VP).

The Chinese word segmentation separates a maritime affair record into three different paragraphs and extracts necessary information for ED, SID and PD. The algorithm for the Chinese word segmentation is shown in Table 1. It relies on a domain dictionary (*DicBase*) and a text tree (*TextTree*). The *DicBase* contains the main vocabulary for a collision avoidance problem. It is created based on the following principles: (1) The vocabulary is arranged in the sequence of WordKindSet= <noun, verb, adjective, adverb, conjunction, pronoun, preposition, auxiliary, quantifier, numeral>; (2) The words having same initial character are arranged in length sequence from long to short; (3) Each word in the vocabulary is appended a numerical value to express its occurrence frequency. A higher frequency is associated with a larger value. The *TextTree* is generated by paragraph segmentation, sentence segmentation and node segmentation. The three paragraphs corresponding to the ED, SID, and PD respectively are segmented. The sentence segmentation is based on the original text sentence and the segmentation tag is the five punctuations, "。", "，", "：", "？" and "！". The node segmentation is based on the kind of the character strings. There are five kinds of characters which are considered, NodeKindSet = <Chinese, English, number, special symbol, quotation mark>. Only the Chinese need to be segmented furthermore, and the other four kinds of character strings can be treated as a solid semantic unit and do not need more treatment.

Table 1: Algorithm for Automatic Segmentation

```
INPUT:     TextTree, DicBase;
OUTPUT:    SegBase;
INITIALIZATION: NodeCount=0; DicCount=0; SegCount=0; SegFlg=FALSE; DelWord ←NULL;
           RemainWord ←NULL;
BEGIN:
   WHILE (TextTree[NodeCount] is not NULL) DO
   {  InBuffer ← TextTree[NodeCount]; NodeCount++;
      IF InBuffer.Kind ∈{English, Number, Symbol and quotation }
     ELSE {   WHILE ( Infuffer.Words is not NULL) DO
        {  DicCount=0; SegFlg=FALSE;
             WHILE (DicBase[DicCount] is not NULL]) DO
             {  IF (InBuffer.Words == DicBase[DicCount])
                THEN {   SegBase[SegCount] ← InBuffer.;
                         SegCount++; SegFlg=TRUE;
                         nBuffer ← DelWord;
                } ELSE DicCount ++;
              }
             IF (SegFlg==FALSE)   THEN {
                FMaxMatch: DelWordPro (LastOne, InBuffer, RemainWord, DelWord);
                BMaxMatch: DelWordPro (FirstOne, InBuffer, RemainWord, DelWord);
                 IF (RemainWord is NULL) THEN
                 {  SegBase[SegCount] ← InBuffer.;
                    SegBase[SegCount].Kind= unknown;
                    SegCount++;
                  } ELSE InBuffer ← RemainWord;
                }
             }
          }
     }
END
```

This is an FMaxMatch and BMaxMatch algorithm. It runs in both forward and backward directions using the final word list as the references. Some domain knowledge is used in the algorithm to improve the segmentation efficiency. The outputs of FMaxMatch and BMaxMatch are stored in a static database, *SegBase*. The differences between the FMaxMatch and BMaxMatch outputs indicate the positions where the overlapping ambiguities occur. To avoid the ambiguity in segmentation, three rules are used to remove the ambiguity in our algorithm. The first rule is to remove overlapping ambiguity. The algorithm detects it and dispels it by selecting the words with higher occurrence frequency as the segmentation result or selecting the words manually. The second rule is to remove combination ambiguity by assigning a high priority to a combined string. The third rule is to deal with an unrecorded ambiguity string which is detected as *SegBase.Kind.* For an unrecorded ambiguity string, the "unknown" will be assigned to *SegBase.Kind.*

B. Creating a template case

Having P_c and S_c from the previous process, this process creates a potential case, $c_{tmp} = \{p_{ctmp}, s_{ctmp}, m_{ctmp}\}$ (where $p_{ctmp} \equiv p_c, s_{ctmp} \equiv s_c; m_{ctmp}$ is to be determined). A potential case is a structured case representation, which might be added to a case base as a new case or be merged with the other cases based on the case base maintenance policies. These policies are presented in the following section.

C. Updating the case base

In ship navigation, an encounter situation (collision case) might occur many times, and the avoiding action may be either the same or different. In such a case, we expect to create a single case to restore these experiences rather than multiple cases. Therefore, we need a sophisticated approach to manage the case base when we add a potential case to an existing case base. The main goal of this algorithm is to determine the attributes of m_c for a given temporary case. The first step is to determine whether the potential case could be a new case. We check the redundancy or inconsistency of the potential case against the existing case base. If a case is not against any case in the existing case base, this case could be a new case. We add it to the case base. Otherwise, we move on to the second step that conducts case base management for the existing case base if we find a case (c_i) that is similar to c_{tmp}. This includes updating an existing case in the case base, deleting a case, and merging multiple cases into a new case. This operation is realized by updating the attributes for m_c. If we detected a similar case (c_i) in the existing case base against the potential case c_{tmp}, i.e., $p_{ci} \approx= p_{ctmp}$[1] and $s_{ci} \approx= s_{ctmp}$, then m_{ci} will be updated to reflect the effect of the repair action applied to the problem. If c_{tmp} is a positive case, then we increase the count of successful repair actions of m_{ci} otherwise we increase the count of unsuccessful repair actions of m_{ci}. In the same way, if we detected a similar case (c_i) against case c_{tmp}, which has similar problem descriptions but different solutions, i.e., $p_{ci} \approx= p_{ctmp}$ and $s_{ci} \approx! s_{ctmp}$, we will update the existing case by adding the new solution to it, so that, the case will become more powerful for solving the similar problem in the future.

---

[1] $\approx=$ means that two items are similar. It is computed with the nearest neighbor algorithm in our system.

4. Experiments

We implemented a case-learning component using the proposed method with the support of the CBR engine in our developed multi-agent system in a VC++ platform. We conducted some experiments for creating cases from the collected maritime affair records.

We have collected 60 collision avoidance cases from maritime affairs record books [22, 23]. These records were written in Chinese and cover a time from 1972 to 2006. Most of the records contain full information in unstructured Chinese text format. From the collision records, we could extract necessary information for creating cases: including, ES (EI, BI, RI, PA), VP (OS, TS, T, ES), and actions taken for collision avoidance, or reason analysis for collisions. Among 60 cases, 50 collision cases took place in Europe and were collected in Lloyd's Report, and 10 cases were from China. Most of collision or collision avoidance took place in near coast and shallow water area. From the encounter situation, most of the cases are two ship collisions. Only five cases are related to multiple ship collisions. From the viewpoint of encounter relationships, 12 cases are heading collisions; 26 cases are crossing collisions; 10 cases are overtaking collisions; and 10 cases are out navigating route in shallow water area. From the viewpoint of navigation environments, 27 cases happened in an invisible weather; and 33 cases were under visible weather. We created electronic version for these records in the Chinese text format. We provided these electronic documents to our developed case learning system and created cases automatically. At the end, we generated 48 cases successfully. Some cases were created from several collision cases because those cases might contain similar encountering situation and navigating environment and took the similar action for avoiding the collisions. The interesting fact is that only two collision cases do not contain all necessary information for generating the cases.

5. Conclusion and Future Work

In this paper, we started from an overview on the developed CBR-based agent system for collision avoidance. We introduced a method for automatic case generation, which was developed using techniques from Chinese Language Processing and CBR. Even though the proposed method is built on Chinese language processing techniques, it is easy to move on to English or other western language. We also presented the experiments conducted for case learning from the collected 60 collision cases along with some results. The experimental results show that the proposed method can provide a useful and effective means for case creation in CBR-based collision avoidance systems.

Although the proposed method can be used to generate the valuable cases from historic maritime affair records, the created cases have to be evaluated carefully before they are applied to CBR-based collision avoidance systems. As mentioned, we have incorporated the evaluation tool into the multi-agent systems. This tool is capable of evaluating the ship-handling results by analyzing the trajectories. Therefore, we can evaluate the case by conducting ship-handling simulation with the created cases. From the ship-handling results we can evaluate the case quality by analyzing the collision avoidance trajectories. Using ship-handling simulation is a cost effective way for evaluating the cases. Some work is ongoing; we will report the results in other paper. Furthermore, some uncertainties in case presentation and case learning still need to be studied in future.

Acknowledgments

This research is supported by Shanghai Leading Academic Discipline Project (S30602) and Shanghai Maritime University project (2009130). Part of work was done while Dr. Yuhong Liu visited the National Research Council Canada in 2007.

References

1. Y. Sato and H. Ishii, "Study of collision-avoidance system for ships," *Control Engineering Practice*, vol. 6, pp. 1141-1149, 1998.
2. Y. Liu and W. Yang, "The structure Design of an Intelligent Decision Support System for Navigation Collision Avoidance," *IEEE The Third International Conference on Machine Learning and Cybernetics*, vol. 1, pp 302-307, August 2004.
3. Y. Liu, "A design and Study on Intelligence Collision Prevention Expert System for Navigation," *Ph.D. thesis, Harbin Engineering University*, *china*, July 1999.
4. C. Yang, "An Expert System for Collision Avoidance and Its Application," *Ph.D. thesis, Hiroshima University, Japan,* September 1995.
5. C. Hwang, "The Integrated Design of Fuzzy Collision Avoidance and H_∞- Autopilots on ships," *The Journal of navigation*, vol. 55, no.1, pp.117-136, 2002.
6. Y. Liu, X. Du, and S. Yang, "The Design of a Fuzzy-Neural Network for Ship Collision Avoidance," *ICMLC 2005*, *Lecture Notes in Computer Science,* vol. 3930, pp. 804-812, July 2006.
7. I. Watson and F. Marir, "Case-Base Reasoning: A Review", *Knowledge Engineering Review*, vol. 9, no. 4, 1994.
8. C. Yang, B. Orchard, B. Farley, and M. Zaluski, "Automated Case Base Creation and Management," *Proceedings of International Conference on Industrial & Engineering. Applications of Artificial Intelligence & Expert System (IEA/AIE-2003)*, June 2003.
9. M. Therani, J. Zhao, and B. Marshall, "A case-based reasoning framework for workflow model management," *Data & Knowledge Engineering,* vol. 50, pp. 87-115, 2004.
10. P. Avesani, S. Ferrari, and A. Susi, "Case-Based Ranking for Decision Support Systems," *ICCBR 2003*, LNAI 2689, PP. 35-49, 2003.
11. D. McSherry, "Similarity and Compromise," *ICCBR 2003, LNAI 2689*, pp. 291-305, 2003.
12. J. Nordlund and H. Schafer, "Case-Based Reasoning in a support System," Master's thesis, UMEA University, Sweden, April 2006.
13. Y. Liu and S. Hu, "An Evaluation System for Single-Target Ship Collision Avoidance Based on Data Fusion," *NAVIGATION OF CHINA*, vol. 65, no. 4, pp. 40-45, Dec. 2005.
14. Y. Liu and H. Liu, "Case Learning Based on Evaluation System for Vessel Collision Avoidance," *The Proceedings of IEEE The Fifth International Conference on Machine Learning and Cybernetics*, vol. 4, pp. 2064-2069, August 2006
15. Y. Liu, M. Wen, and Z. Du, "A case Learning Model for Ship Collision Avoidance Based on Automatic Text Analysis," *The Proceedings of IEEE The Fifth International Conference on Machine Learning and Cybernetics,* July, 2009, Baoding, China.
16. C. Yang, B. Farley, and B. Orchard "Automated Case Creation and Management for Diagnostic CBR Systems", *International Journal of Applied intelligence*, Vol. 28, No. 1, 2008, pp. 17-28
17. Y. Liu, C. Yang, and X. Du," Multi-agent Planning for Ship Collision Avoidance". *The Proceedings of IEEE International Conferences on Cybernetics & Intelligent Systems (CIS) and Robotics, Automation & Mechatronics (RAM) (CIS-RAM 2008),* Chengdu, China, June, 2008.
18. Y. Liu, C. Yang, and X. Du, "A CBR-Based Approach for Ship Collision Avoidance ", *The Proceedings of the 21st International Conference on Industrial & Engineering Applications of Artificial Intelligence & Expert Systems (IEA/AIE-2008),* Wroclaw, Poland, June 2008
19. A. S. Rao and M. P. Georgeff. Modeling Rational Agents within a BDI-Architecture. *In Proceedings of the 2nd International Conference on Principles of Knowledge Representation and Reasoning*, pages 473-484, 1991
20. H. Gong, C. Gong, and C. Zhou, "Chinese word segmentation system research", *Journal of Beijing Institute of Machinery*, Vol. 19, No. 3, pp.52-61, Sep. 2004.
21. J. Qiu, T. Wen, and L. Zhou, "Research of Chinese Automatic Segmentation and Content Analysis Method", *Journal of The China Society For Scientific and Technical Information*, Vol. 24, No. 3, pp.309-317, June. 2005
22. China MSA (Maritime Safety Administration). "Typical water traffic accident cases", China Communications Press, pp.1-89, Oct. 2007.
23. J. Zhao and F. Wang. "The collision regulations and cases", Dalian Maritime University Press, pp.560-685, Mar.1997.