Applying Case-Based Reasoning Technology in Electronic Commerce Environment

Nyein Ei San, Win Lei Lei Phyu
University of Computer Studies, Yangon
tonetone.nyein@gmail.com,winlei@gmail.com

Abstract

Case-based reasoning technology opens a very promising and enriching field for electronic commerce application, medical diagnosis, knowledge management, process control and so on. Case-based reasoning (CBR) is an approach to problem solving and learning, by reusing the solution to similar problem stored as cases in a case base. Existing electronic shops based on Case-based Reasoning technology allow customers to search for adequate products by only specifying the attributes for the products. They do not further support customers after the retrieval step. In this system, Case-based reasoning technology is applied for electronic commerce application on the web. This approach is not only to find the most similar product but also to offer adaptation phase after retrieval step.

Keywords: Case-Based reasoning, Electronic Commerce, Case Retrieving, Adaptation

1. Introduction

Electronic commerce offers a huge variety of application areas nowadays. Much effort is put into research and exploration of these areas. However, research to address one of the most important problems, responsive online sales support, has generated only few promising approaches [5]. Most electronic catalogs and online shops do not explore the interactively available on the Web [6].

Driven by the challenge to improve Internet-based electronic commerce, case-based reasoning (CBR) technology emerged as a new important possibility [6].

The main role for CBR in electronic commerce is intelligent sales support, i.e., the task to select a product from a list of alternative which is most appropriate for the demands of the customer. However, most of the electronic commerce application using CBR can only support retrieval step. Customer entering an electronic shop may have certain specific ideas about the product they are looking for. They specify their ideas, requirements, or wishes in a query. It will very likely happen that the products found do not exactly meet the customer's wishes. Thus, the retrieved products have to be adapted to these wish [5].

In this paper, section 2 presents the related work of the system. Section 3 presents background theory of CBR. Section 4 describes nearest neighbor algorithm. Section 5 presents adaptation. Section 6 describes the system implementation and section 7 presents the experimental result. The conclusion is described in section 8.

2. Related Work

Case-Based reasoning can be used in many application areas. I.Vollarth proposed case-based reasoning technology in online sales catalog system. W.Wilke described the use of case-based reasoning technology in electronic commerce.

Furthermore, case-based reasoning is important technology in many applications. Maurice Bopp studied the new CBR-based application for voxelised phantom creation. Fouzia Anguel used cased-based reasoning for fault diagnosis of gas turbines

3. Case-Based Reasoning

3.1. Case-Based Reasoning Technology

A new problem is solved by retrieving one or more previously experienced cases, reusing the case in one way or another, revising the solution based on reusing a previous case, and retaining the new experience by incorporating it into the existing case-base.

A case is placed in a Case Base and is called source case with which one will find out how to solve a new case. This new case is called "target" case. In Electronic Commerce environment, a target case (new problem) is typically the assignment of a particular product to a set of demand (or requirements) stated by the customer [5].

3.2. Case-Based Reasoning Cycle

Aamodt and Plaza (1994) proposed a first CBR cycle to make evident the knowledge engineering effort in CBR. A general CBR cycle can be described by following four processes.

1. RETRIEVE the most similar case or cases;
2. REUSE the information and knowledge in that case to solve the problem;
3. REVISE the proposed solution if necessary, and
4. RETAIN the new solution as a part of a new case for future problem.

The case-based reasoning cycle is illustrated in Figure 1.

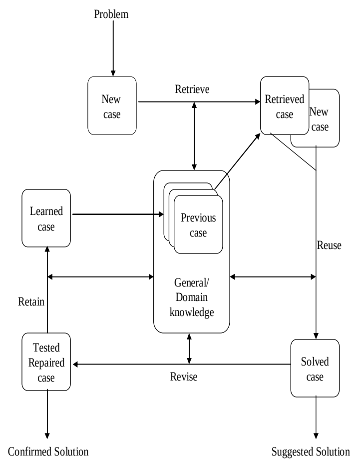

Figure 1: Case-based Reasoning Cycle

3.3. Case Representation

The representation of problem in CBR is primarily the problem of deciding what to store in case, finding an appropriate structure for describing case contents, and deciding how the case memory should be organized and index for effective retrieval and reuse. In this system, Attribute value representation method is used for case representation. Problem description consists of a set of attributes and values. That attribute predict a solution [1].

4. Nearest Neighbor Algorithm

The K-nearest neighbor approach is the one way to define the similarity between new case and cases already stored in the case base.
The similarity values are ordinal values that are often normalized to interval [0, 1]. A value of "0" means "does not satisfy the query at all" and a value of "1" say "that's exactly what you asked for" [5].

The useful approach to define the similarity measure is to start with "local similarity measure" for the all attributes [5].
The similarity measure for all attribute can be calculated as follow:

$$\mathrm{Sim}(q_i, c_i) = 0 \qquad \text{if } q_i \neq c_i$$
$$\mathrm{Sim}(q_i, c_i) = 1 \qquad \text{if } q_i = c_i$$

Where,

q is the target case (query case?)

c is the source case (case in the case library).

Once the local similarity functions are defined, the global similarity of two cases must be derived from the local similarities. The usual way to do this is to apply a weighted sum to all the local similarities, [3]. The similarity between q and c can be calculated from the local similarities as follows:

$$\mathrm{Sim}(q, c) = \frac{\sum_{i=1}^{n} w_i \, \mathrm{Sim}(q_i, c_i)}{\sum_{i=1}^{n} w_i}$$

Where,

q is the target case.

c is the source case.

n is the number of attributes in each case.

i is the individual attribute from 1 to n.

w_i is the importance weighting of attribute i.

Thus, the weight is introduced in the case retrieval and the similarity between cases is considered to be weighted summation of the similarity between attributes.

4.1. Assigning weight value to case feature

Feature weight for most problem domains are content dependent. The weight assign to each feature of the case tells how much attention to pay to matches and mismatches in the field when computing the distance measure of a case. There are two ways of assigning weight value to case features.

(1) One way to assign importance value is to have a human expert assign them as the case library is being built.

(2) Another way is to assign to important values to do a statistical evaluation of a know corpus of cases to determine which dimension predict different outcomes and/or solution best.

In this system, the first method is used to define the weight value.

5. Adaptation

Some CBR system do not only support the retrieval process but they also adapt case solutions to the new problem and thus create new solutions. Of course, this requires the solution, here the product, to be configurable.

The knowledge needed to perform this kind of adaptation must be represented in some suitable form during the development of the CBR system. Depending on the application domain, the adaptation process can be more or less complicated. One possible representing is a set of rules that perform certain actions given that the required preconditions for the actions are valid [5].

6. System Implementation

This system is to use the case based reasoning technology for product selection in Electronic commerce, for Computer Shop. It uses the components of the computer that the user wanted and compares these with the components of the case already saved in its case base.

The system retrieves the best matches, which are most similar to the requirement. If the results are not satisfied, these components can be changed by using adaptation rule. If the system can adapt the component, system shows the adapted product and if not, shows the message why cannot adapt. The system flow diagram is illustrated in Fig 2.

7. Experimental Result

In this section, the computer shop is used as an example of electronic commerce system. This

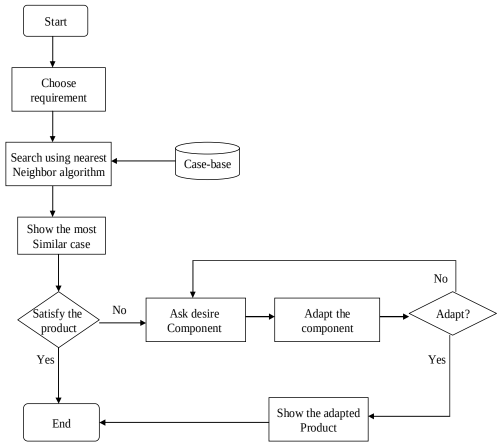

Figure 2: Data flow diagram of the system

system uses 1000 cases and this data are received from Direct Channel Distribution Computer sales and service. In this system, attribute of components are represented in case by attribute value representation. There are twelve attributes for component of the computer. This attributes and their values are represented in Table 1.

Table 1: Attribute value table for component of computer

Compo-nent	Values
Brand	1 = Compaq , 2 = Dell, 3 = Acer , 4 = Asus, 5= Samsung, 6 = HP, 7 = Prolink, 8 = ViewSonic
CPU	1 = Intel Celeron 1.8 GHz 2 = Intel Dual Core 2.2 GHz 3 = Intel Dual Core 2.5 GHz 4 = Intel Core 2 Duo 2.8 GHz 5 = Intel Core 2 Duo 2.93 GHz 6 = Intel Core 2 Quad 2.66 GHz 7 = Intel Core 2 Quad 2.5 GHz 8 = Intel Core 2 Quad 3.0 GHz
Memory	1 = 1GB DDR2 800MHz 2 = 2GB DDR2 800MHz 3 = 1GB DDR3 1066MHz 4 = 2GB DDR3 1066MHz 5 = 4GB DDR3 1066MHz
Hard disk	1 = 80, 2 = 160, 3 = 250, 4 = 320, 5 = 500
Graphic card	1 = PCIE (512MB) AH3450 ASUS 2 = PCIE (512MB) 8400GS MSI 3 = PCIE (1GB) N9500GT GIGABYTE 4 = PCIE (1GB) N9500GT MSI 5 = PCIE (1GB) N9500GT ASUS
Screen size	1 = 13.0, 2 14.0, 3 = 15.0, 4 = 15.6, 5 = 16.0, 6= 17.0 , 7 = 18.5, 8 = 19.0, 9 = 20.0, 10 = 21.0, 11 = 22.0, 12 = 23.0

Writer	1 = Sony DVD ROM, 2 = Sony DVD R/W 3 = AOPEN DVD R/W, 4 = HP DVD R/W 5 = LITEON DVD R/W
Mouse	1 = (Blaze) Optical Mouse (PS/2) 2 = (Blaze) Optical Mouse (USB) 3 = (Sony) Optical Mouse (USB) 4 = (Prolink) Optical Mouse (PS/2) 5 = (Prolink) Optical Mouse (USB) 6 = (A4Tech) Optical Mouse (PS/2) 7 = (A4Tech) Optical Mouse (USB)
Key-board	1 = A4Tech (PS/2), 2 = A4Tech (USB) 3 = Prolink (PS/2), 4 = Prolink (USB) 5 = None
Mother-board	1 =MSI _P4 G31M3L 2 =MSI _P4 P31 NeoF-V2 3 =MSI _P4 P31 Neo3-F 4 =ASUS P4- P5- KPL- AM 5 =ASUS P4- P5- KPL CM 6 =ASUS P4- P5- KPL/1600 7 =ASUS P4- P5P41T-LE 8 =GIGABYTE P4 G31M ES2C 9 =GIGABYE P4-P43-ES3G 10 =GIGABYTE P4-EP43-DS3L 11 =Acer Aspire e360 12 =Acer Aspire m1640 13 =Acer Aspire 9300 14 =Acer Aspire T135 15 =Dell F3542 Inspiron 1150 16 =Dell RJ272 Inspiron B120 17 =Dell Alienware M11x 18 =Compaq EVO D300P4 19 =Compaq 217155-002 20 =HP 172624-001 21 =HP F1580-69001 22 =HP F2157-60937
Laptop	1 = TRUE , 2 = FALSE
Cost	1 = $ 300, 2 = $ 400, 3 = $ 500, 4 = $ 600 5 = $ 700, 6 = $ 800, 7 = $ 900, 8 = $ 100

To retrieve the most similar case, the system calculates the similarity between old cases and new case. Calculating similarity measure needs weight value for each attribute. The weight value, which is defined by expert, is illustrated in Table 2.

Table 2: Weight value table for attribute

No	Component	Weight value
1	Brand	1
2	CPU	0.8
3	Memory	0.7
4	Hard disk	1
5	Graphic card	0.8
6	Screen size	0.4
7	Writer	0.6
8	Mouse	0.2
9	Keyboard	0.2
10	Motherboard	1
11	Laptop	1
12	Cost	1

Table 3 illustrates the attribute value of the two old cases and new case.

Table 3: The example of two old case and new case

No	Component	Old case 1	Old case 2	New Case
1	Brand	1	1	1
2	CPU	3	4	3
3	Memory	2	2	2
4	Hard disk	4	4	3
5	Graphic card	4	3	4
6	Screen size	7	9	7
7	Writer	2	2	2
8	Mouse	4	1	4
9	Keyboard	3	1	1
10	Motherboard	3	4	3
11	Laptop	2	2	2
12	Cost	4	4	4

The following example shows how to calculate the similarity between the new case and old cases by using nearest neighbor algorithm.

For old case 1;

Sim $(q_1, c_1) = 1$ Sim $(q_7, c_7) = 1$
Sim $(q_2, c_2) = 1$ Sim $(q_8, c_8) = 1$
Sim $(q_3, c_3) = 1$ Sim $(q_9, c_9) = 1$
Sim $(q_4, c_4) = 0$ Sim $(q_{10}, c_{10}) = 0$
Sim $(q_5, c_5) = 1$ Sim $(q_{11}, c_{11}) = 1$
Sim $(q_6, c_6) = 1$ Sim $(q_{12}, c_{12}) = 1$

$$\frac{\sum_{i=1}^{n} w_i \, \mathrm{Sim}(q_i, c_i)}{\sum_{i=1}^{n} w_i} = 0.86$$

For old case 2;

Sim $(q_1, c_1) = 1$ Sim $(q_7, c_7) = 1$
Sim $(q_2, c_2) = 0$ Sim $(q_8, c_8) = 0$
Sim $(q_3, c_3) = 1$ Sim $(q_9, c_9) = 1$
Sim $(q_4, c_4) = 0$ Sim $(q_{10}, c_{10}) = 0$
Sim $(q_5, c_5) = 0$ Sim $(q_{11}, c_{11}) = 1$
Sim $(q_6, c_6) = 0$ Sim $(q_{12}, c_{12}) = 1$

$$\frac{\sum_{i=1}^{n} w_i \, \mathrm{Sim}(q_i, c_i)}{\sum_{i=1}^{n} w_i} = 0.43$$

In this above calculation, the new problem is close to the case 1 than case 2 because case 1’s similarity value is more than case 2’s similarity value. Therefore, the case 1 is retrieved and the solution is ‘found the most similar case'. A set of computer and its cost is presented to the user.

After retrieving, the retrieved products are not satisfied, this product can be adapted.

In this system, the adaptation result can have three states. These are:

(1) the product can be change
(2) the product can be change but the price exceed the user need
(3) the product cannot be change

The example of adaptation is described in table 4.

Table 4: The example of adaptation table

Compaq		
Intel Dual Core 2.5 GHz	Change CPU	Intel Dual Core 2.8 GHz
2GB DDR2 800MHz		
320 GB		
PCIE (1GB) N9500GT MSI		
18.5”		
Sony DVD R/W	change writer	HP DVD R/W
(Prolink) Optical Mouse (PS/2)		
Prolink (PS/2)		
MSI _P4 P31 Neo3-F	Change motherboard	ASUS P4-P5P41T-LE
False		
566		

Table 5 shows the adaptation components and adaptation result.

Table 5: The adaptation component and its result

	Adapted component	Adaptation Result
Example 1:	If CPU was changed	It can be change, but exceed the price (state 2)
Example 2:	If writer was changed	It can be change (state 1)
Example 3:	If motherboard was changed	It cannot be change (state 3)

8. Conclusion

In this system, case-based reasoning technology is applied in electronic commerce environment. The computer shop is used as an example of Electronic commerce application. These systems retrieve the computer set that are most similar to requirement. Moreover, this system offers adaptation after retrieval step. Therefore, this system can be used conveniently on the web at any place and any time.

9. References

[1] A.Aamodt & E.Plaza (1994). Case-Based Reasoning: Foundational Issues, Methodological Variations, in AICOM, Vol. 7, No. 1, 1994, pp. 39-59.

[2] Fouzia Anguel, Mokhtar Sellami. Knowledge Management for Fault Diagnosis of Gas Turbines Using Case Based Reasoning. Badji Mokhtar Annaba University.

[3] I.Vollarth, W.Wilke, & R.Bergmann (1998). Case-based reasoning support for online catalog sales. IEEE Internet Computing, July-August 1998.

[4] Maurice Bopp, Julien Henriet, Brigitte Chebel-Morello, Libor Makovicka, David Broggio. Preliminary study of a new CBR-based application for voxelised phantom creation: REEPH.

[5] Sascha Schmitt, Ralph Bergmann (1999). Applying Case-Based Reasoning Technology for Product Selection and Customization in Electronic Commerce Environments, Twelfth International Bled Electronic Commerce Conference, Slovenia, June 7- 9, 1999.

[6] U.J.Timm, & M.Rosewitz (1998). Electronic Sales Assistance for product Configuration. Proceedings of the Eleventh International Bled Electronic Commerce Conference, Volume 1- Research, 1998.

[7] W.Wilke, M.Lenz, & S.Wess (1998). Case-Based Reasoning for Electronic Commerce. In Lenz et al. (Eds.): Case-Based Reasoning Technology from Foundations to Applications, Springer-Verlag.