

Case-Based Reasoning

Watson chapters 1 - 4

Bill Cheetham, Kai Goebel

Slides modified from Dr. Ian Watson

1

Contents

- What is Case-Based Reasoning?
- How does CBR work?
- Advantages / Disadvantages
- Who Uses CBR?
 - case study - Lockheed
 - lists at web sites
 - GE



2

What is CBR?

- A case-based reasoner solves new problems by using or adapting solutions that were used to solve old problems
- offers a reasoning paradigm that is similar to the way many people routinely solve problems

3

What Is CBR?

- How will you get home?
- Generate a path
- Or remember the way
- Path to new location?
- Remember close path
- Adapt it



What Is CBR?

What is 12×12 ?

144

What is 12×13 ?

$12 \times 12 + 12$

156



What Is CBR?

In order to have a phone in every house 1/10 of the entire US population would need to be phone operators

Create a phone book with people and their phone numbers



6

Who uses CBR?

Lawyers

- find previous ruling that applies to case
- show that it applies to current case

Real Estate Appraiser

- find similar comparable houses
- estimate value of target based on value of comparable

7

Quotes

- "We know nothing of what will happen in future, but by the analogy of experience." -- Abraham Lincoln
- "Study the past if you would divine the future." -- Confucius
- "If at first you don't succeed, you are running about average." -- Bill Cosby

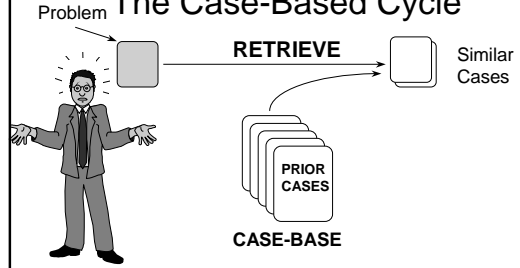
8

The Case-Based Cycle



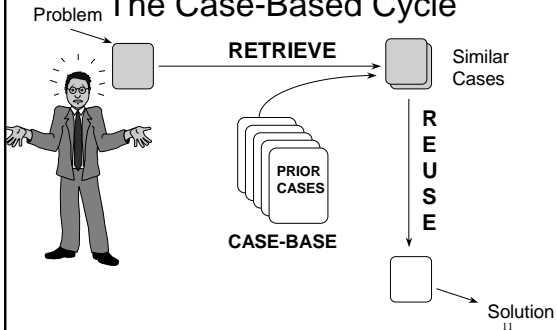
9

The Case-Based Cycle



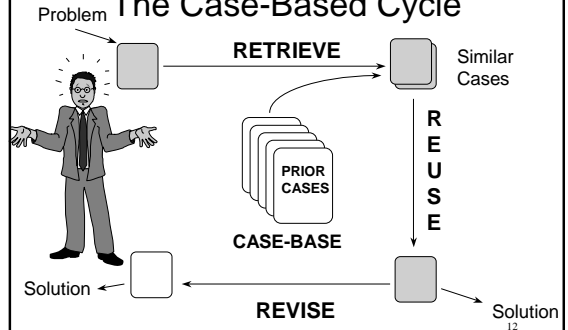
10

The Case-Based Cycle

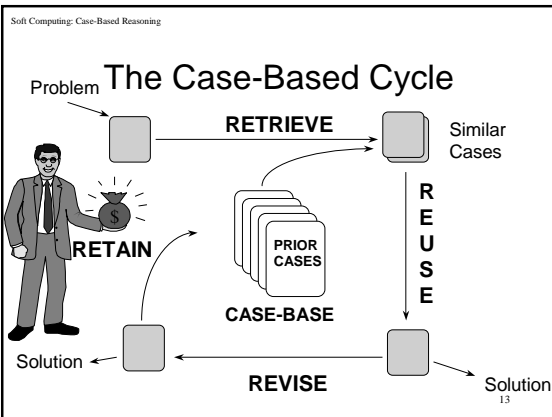


11

The Case-Based Cycle



12



Soft Computing: Case-Based Reasoning

History of CBR

CBR can trace its roots to the field of psychology and theories about how human memory works

- "Episodic Memory" [Tulving 1972] provides a method for storing and recalling large chunks of related information such as events, scenes, occurrences, and stories
- "schema" [Rumelhart 1977] reasoning is the process of applying chunks of information to new situations

14

Soft Computing: Case-Based Reasoning

History of CBR- scripts

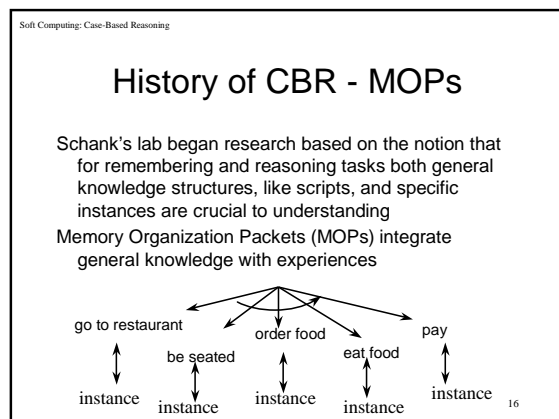
Roger Schank's group invented scripts at Yale natural language lab during the mid-seventies

"scripts represent generalizations about actions that should take place in stereotypical situations"

Restaurant sequence of events:

1. enter restaurant,	7. order food,
2. be seated by hostess,	8. wait for waiter,
3. obtain menu from waiter,	9. waiter returns with food,
4. order drinks,	10. eat food,
5. waiter leaves,	11. waiter returns with bill,
6. waiter returns with drinks,	12. pay bill & leave

15



Soft Computing: Case-Based Reasoning

History of CBR - TOPs

Thematic Organization Packets (TOPs) categorize situations by the plans of the participants rather than the details of the situation

Schank's daughter was diving in the ocean looking for sand dollars. He pointed out where a group of them were, yet she continued to dive elsewhere. He asked why, and she told him that the water was shallower where she was diving. This reminded him of the old joke about a drunk searching for his lost ring under the lamppost where the light was better.

17

Soft Computing: Case-Based Reasoning

History of CBR - Software

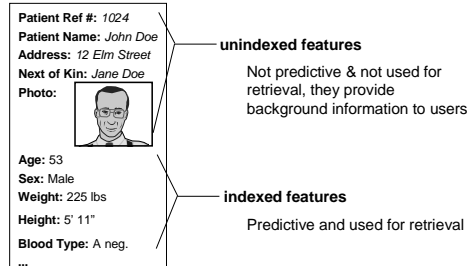
- The 80's
 - the original CBR programs CASEY, CHEF, JULIA were written in LISP
 - research tools in public domain
- The 90's
 - the development of commercial CBR tools, mostly in C
- Today
 - commercial Web-based Java tools
 - specific application (customer self-service)¹⁸

What is a Case?

- several features describing a problem
- plus an outcome or a solution
- cases can be very rich
 - text, numbers, symbols, plans, multimedia,
- cases are not usually distilled knowledge
- cases are records of real events
- and are excellent for justifying decisions

19

2 types of case features



20

What is a Case-Base?

- A case-base is a set of cases.
- Case-bases are usually just flat files or relational databases

a robust case-base, containing a representative and well distributed set of cases, is the foundation for a good CBR system

[Kriegasman & Barletta, 1993].

21

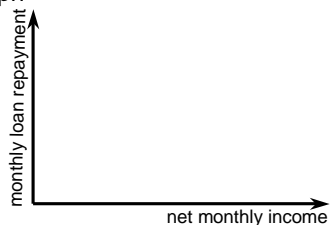
How Does Retrieval Work?

- imagine a decision with two factors that influence it
- should you grant a person a loan?
 - ① net monthly income
 - ② monthly loan repayment
- more factors in reality



How Does CBR Work?

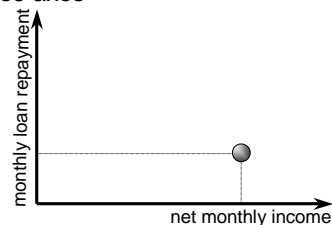
- these factors can be used as axes for a graph



23

How Does CBR Work?

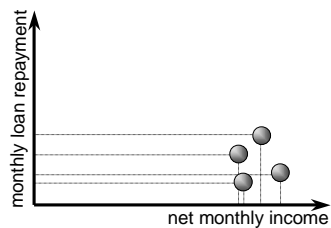
- a previous loan can be plotted against these axes



24

How Does CBR Work?

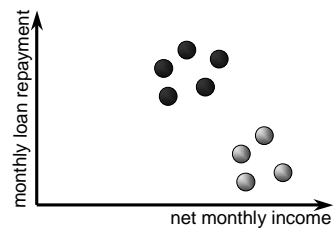
- and more good loans



25

How Does CBR Work?

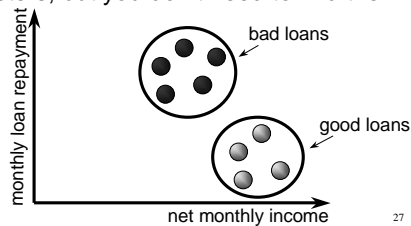
- plus some bad loans



26

Lazy Learning

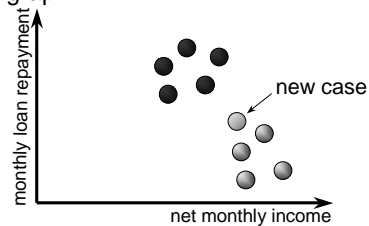
- past cases (loans) may tend to form clusters, but you don't need to find them



27

How Does CBR Work?

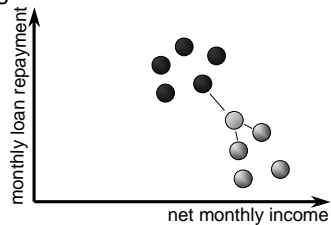
- a new loan prospect can be plotted on the graph



28

How Does CBR Work?

- and the distance to its nearest neighbors calculated



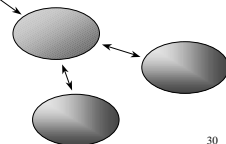
29

How Does CBR Work?

- the best matching past case is the closest



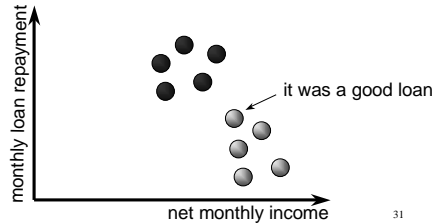
- this suggests a precedent
- the loan should be successful



30

How Does CBR Work?

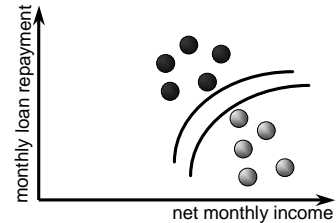
- over time the prediction can be validated



31

How Does CBR Work?

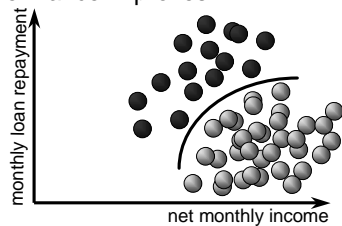
- the system is learning to differentiate good and bad loans better



32

How Does CBR Work?

- as more cases are acquired its performance improves



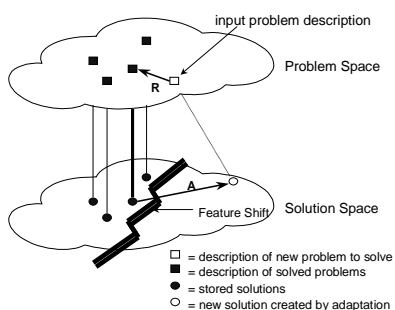
33

Retrieval Issues

- do all indexed features have the same weight?
- Is the similarity linearly proportional to the distance a case is from the new problem?
- What distance measure should be used (city block, line of sight, ...)
- Uniformity of solution space

34

Feature Shift



35

How to weight the features

- First normalize the attributes
 - find min. and max.
 - set min = 0, max = 1
- Real estate appraiser example
 - living-area between 2,000 and 3,500 sq. ft.
 - 2,000 = 0
 - 3,500 = 1

36

How to weight the features

- Ask an expert
- Look for trends in data (plot, regression)
- Use leave-one-out testing
 - select an item from the case base where you know the solution
 - run your CBR system on the case
 - determine error = difference in solution suggested by CBR and actual solution
 - update rules to minimize this error (GA?)

37

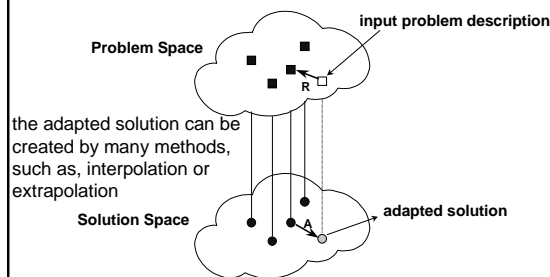
Adaptation

In many situations the case returned is not the exact solution needed

- Many techniques can be used
 - rule-based system
 - heuristic search
 - other
- The best technique depends on the application

38

Adaptation



39

Rule Based Adaptation

Example from real estate appraiser

- retrieved a house selling at \$100,000 that is exactly like the one being appraised except it has a 2 car garage and target has 1 car garage
- the value of an extra garage is \$4,000
- adjusted value is $\$100,000 + \$4,000$

40

How do you learn Adaptations

Analyze the Data

- Look for trends in data (plot, regression)
- Compare similar items
 - Find two data items that are identical except for one difference
 - The solution difference is the value of the change in the attribute
- Evaluate rules by testing different ones

41

Exercise

- Estimate height using students as the case-base
- 5 students come to front and be the case base
- Predict the height of a 6th student

42

Summary

- in real life the problem space is N dimensional
- new features can be added if they become relevant
- feature vectors can be weighted to reflect their relative importance
- tolerant of noise & missing data
- k -Nearest Neighbor Retrieval

43

Advantages of CBR

- CBR is intuitive - it's how we work
- no knowledge elicitation to create rules or methods
- this makes development easier
- systems *learn* by acquiring new cases through use
- this makes maintenance easy
- justification through precedent

44

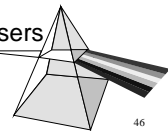
When to Apply CBR?

- when a domain model is difficult or impossible to elicit
- when the system will require constant maintenance
- when records of previously successful solutions exist

45

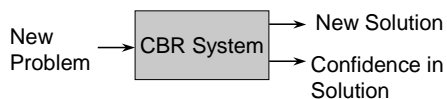
CBR is Transparent

- precedent is an accepted method for justifying a decision
- nearest neighbor retrieves the best matching past cases
- the process is transparent
- i.e., easily understood by users
- this increases acceptance



46

CBR with Confidence

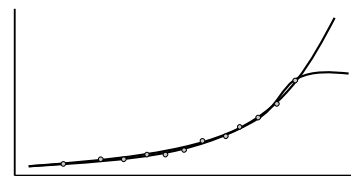


Confidence based on:

- Number of cases matching
- Similarity of matching cases to new problem
- Similarity of matching cases to each other

47

CBR with Confidence



Blue dots - data points

Red line - actual line to be predicted

Yellow line - prediction by a neural network

48

CBR Systems Learn

- decision making is dynamic
- CBR systems learn by acquiring new cases
 - no addition of new rules
 - no retraining of neural networks
 - no re-evolving new populations with new genomes
 - no re-induction of rules from data

49

Case-Base Issues

- How many cases are needed
- How to remove overlapping cases
- How to efficiently search
 - create abstractions from cases
 - multiple case bases
- What features to use for indexing
- How to weight the features

50

Disadvantages of CBR

- Can take large storage space for all the cases
- Can take large processing time to find similar cases in case-base
- Cases may need to be created by hand
- Adaptation may be difficult
- Needs case-base, case selection algorithm, and possibly case-adaptation algorithm

51

Disadvantages of CBR

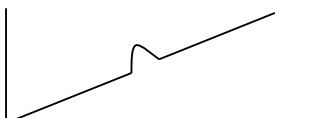
- if you require the *best* solution or the *optimum* solution - CBR may not be for you
- CBR systems generally give *good* or *reasonable* solutions
- this is because the retrieved case often requires adaptation



52

CBR vs linear regression

- linear regression summarizes data while CBR retains all data points
- hard for regression to learn strange shapes



53

CBR vs Rule-Based Systems

- CBR offers a cost-effective solution to the 'knowledge acquisition bottleneck' problem
- CBR systems can learn from experience and so can be self-maintaining
- Rule-based systems are better when it is hard to gather case data

54

CBR vs Rule Based System

- rule-based systems justify decisions by showing a rule trace
- decision grant loan because
rule 24 -> rule 61 -> rule 43 -> rule 202
- rule traces are opaque & can be confusing to users



55

CBR vs NN

- neural nets cannot justify their decisions
- users have to *trust* the computer is always correct
- Neural nets cannot take advantage of domain knowledge



56

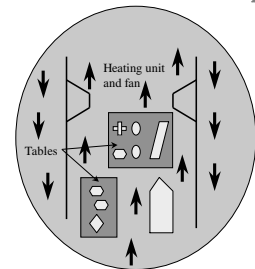
Applications of CBR

- **Classification:** "The patient's ear problems are like this prototypical case of otitis media"
- **Compiling solutions:** "Patient N's heart symptoms can be explained in the same way as previous patient D's"
- **Assessing values:** My house is like the one that sold down the street for \$250,000 but has a better view"
- **Justifying with precedents:** "This Missouri case should be decided just like Roe v. Wade where the court held that a state's limitations on abortion are illegal"
- **Evaluating options:** "If we attack Cuban/Russian missile installations, it would be just like Pearl Harbor"

57

Lockheed - CLAVER

- Lockheed makes aircraft parts from composite materials.
- These materials are made from layers of carbon fibers that are formed into a single component by curing in a large oven, called an autoclave



58

Lockheed



- **PROBLEM** - how to optimize the loading of an autoclave for curing composite materials
- different materials need different heating & cooling procedures
- materials interact with each other in the autoclave
- mistakes are VERY costly

59

Lockheed



- 2 experienced operators relied on plans of previously successful layouts
- New layouts were adapted from old
- If successful they were added to a library
- they wanted to develop a decision support tool to assist experts and to retain expertise as a corporate asset

60

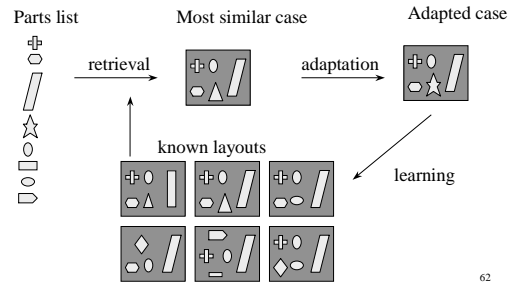
Lockheed



- Lockheed had NO model of the autoclave (a rule-based tool failed)
- the manufacturers could not provide one
- layouts did not repeat
- materials were constantly changing
- designs constantly change
- elements interact
- CBR was used



Lockheed



62

Lockheed



- their system was implemented in 1990
- CLAVIER started with 20 successful layouts
- CLAVIER now has hundreds of successful layouts
- it retrieves a successful layout or adapts one 90% of the time
- acts as a corporate memory

63

CBR Web sites

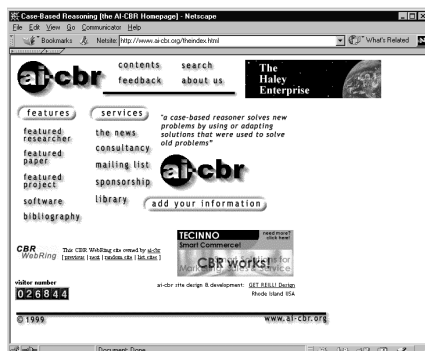
www.ai-cbr.org

- CBR research, places, people, papers, conferences summaries

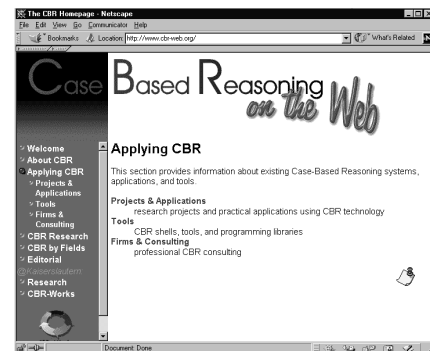
www.cbr-web.org

- German perspective on CBR

64



65



66

