Case Based Reasoning
Case Based Reasoning

An Example

<table>
<thead>
<tr>
<th>Tid</th>
<th>Refund</th>
<th>Marital Status</th>
<th>Taxable Income</th>
<th>Cheat</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Yes</td>
<td>Single</td>
<td>125K</td>
<td>No</td>
</tr>
<tr>
<td>2</td>
<td>No</td>
<td>Married</td>
<td>100K</td>
<td>No</td>
</tr>
<tr>
<td>3</td>
<td>No</td>
<td>Single</td>
<td>70K</td>
<td>No</td>
</tr>
<tr>
<td>4</td>
<td>Yes</td>
<td>Married</td>
<td>120K</td>
<td>No</td>
</tr>
<tr>
<td>5</td>
<td>No</td>
<td>Divorced</td>
<td>95K</td>
<td>Yes</td>
</tr>
<tr>
<td>6</td>
<td>No</td>
<td>Married</td>
<td>60K</td>
<td>No</td>
</tr>
<tr>
<td>7</td>
<td>Yes</td>
<td>Divorced</td>
<td>220K</td>
<td>No</td>
</tr>
<tr>
<td>8</td>
<td>No</td>
<td>Single</td>
<td>85K</td>
<td>Yes</td>
</tr>
<tr>
<td>9</td>
<td>No</td>
<td>Married</td>
<td>75K</td>
<td>No</td>
</tr>
<tr>
<td>10</td>
<td>No</td>
<td>Single</td>
<td>90K</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Refund | Marital Status | Taxable Income | Cheat
---|----------------|----------------|-------
No       | Single         | 75K            | ?     |
Yes      | Married        | 50K            | ?     |
No       | Married        | 150K           | ?     |
Yes      | Divorced       | 90K            | ?     |
No       | Single         | 40K            | ?     |
No       | Married        | 80K            | ?     |

Experts:
- Rules: data mining
- Cases: CBR
Objectives

Show how CBR works.

To introduce the basic components of CBR systems.

Demonstrate some examples of CBR.
Case Based Reasoning

What is CBR?

• Reasoning that adapts previous solutions for similar problem in solving new problem in hand
  • Many problem decision makers encountered are similar to old cases
  • Often more efficient to start with the previous solution to a similar problem than to generate the entire solution again from scratch
  • Experts solve problem based on previous cases
    • Court legal cases, etc
Case

- 4 Re’s
  - Retrieve
  - Reuse
  - Revise
  - Retain

Aamodt and Plaza, 1994
What is CBR?

- Case Based Reasoning
- Case representation
- Indexing of cases
- New problem
- Case retrieval
- Relevant case(s)
- Case adaptation
- Similarity measure
- Heuristics/rules for adaptation
- Solution evaluation
- Not good
- Case base organisation
- Insert case
- Adaptation

Case Base

- Organisation

Flowchart:
- Case base
  - Indexing of cases
  - New problem
  - Case retrieval
  - Relevant case(s)
  - Case adaptation
  - Similarity measure
  - Heuristics/rules for adaptation
  - Solution evaluation
  - Not good
- Insert case
- Adaptation
Case Based Reasoning

Components of CBR

- **Case representation**
  - **Problem**: describes the state of the world when the case occurred
  - **Solution**: states the derived solution to that problem, and/or
  - **Outcome**: the state of the world after the case occurred

- Text, numbers, symbols, plans, multimedia
- Usually (attribute, value) pairs
Components of CBR

- **Case representation**
  - What to store in a case
    - Appropriate *structure* to describe case contents
  - How to organise and index for effective retrieval and reuse
    - Functionality and ease of acquisition
Case Based Reasoning

Components of CBR

- **Case indexing**
  - Assign indices to cases to facilitate their retrieval
  - Features and dimensions tend to be predictive
  - The system has to retrieve the right case at the right time
  - Predictive, useful, abstract and concrete

- Abstract enough to allow for widening the future use of the case-base;
- Not too abstract to avoid retrieving too many cases
Case Based Reasoning

Components of CBR

- **Case base organisation**
  - Flat memory
    - sequentially in a simple list, array or file
  - Hierarchical organisation
    - large case base
    - only small subset needs to be considered during the retrieval
    - organise specific cases which share similar attributes under a more general structure
Case Based Reasoning

Components of CBR

- **Case base organisation**
  - Flat memory
    - Nearest neighbour
    - Weighting: by experts
  - Hierarchical organisation
    - Tree search
    - Find the node that best matches the input

\[
\sum_{n=1}^{N} w_n \text{sim}(f_n^I, f_n^R)
\]

\[
\sum_{n=1}^{N} w_n
\]
Case Based Reasoning

Components of CBR

- **Case adaptation**
  - Structural adaptation
    - adaptation rules are applied directly to the solution stored in cases
  - Derivational adaptation
    - reuses the algorithms, methods or rules that generated the original solution to produce a new solution to the current problem
  - Simple or complex techniques, depend on the problem domain
Case Based Reasoning

Development of CBR

• **Case representation**
  • Attributes that identify problems
  • Indices for storage and retrieval

• **Similarity measure**
  • Features that explain solutions

• **Adaptation**
  • Domain theory of impact of attributes on solutions

• **Case base organization**
  • A CBR system is heavily dependent on structure and content of case base
## Case Based Reasoning

<table>
<thead>
<tr>
<th>Rule based system</th>
<th>Case based reasoning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rules</td>
<td>Examples, stories</td>
</tr>
<tr>
<td>Difficult to convert knowledge to rules, i.e. difficult to explain rules</td>
<td>Easier to tell stories, i.e. handles exceptions/novel cases in weak domains such as law</td>
</tr>
<tr>
<td>Failure reported when no rules are matched</td>
<td>Learning from both successes and failures</td>
</tr>
<tr>
<td>Difficult to justify the solution</td>
<td>Explanation becomes easier and pervasive</td>
</tr>
<tr>
<td>Easier to validate</td>
<td>Difficult validation</td>
</tr>
</tbody>
</table>
CBR – A modified example

- **Residential property valuation***
  - To determine an estimated value at a given location and given time
  - most common methods used by the human appraisers is to find the recent sales that are comparable with the subject property

- **Case: (attribute, value) pairs**
  - Sale Price £185,000
  - Address Wollaton Road, Nottingham
  - Living Area 2000 sq. ft
  - Lot size 20000 sq. ft
  - Bedrooms 3
  - Bathrooms 2.5
  - ...

* Adapted from (Cheetham et al. 2004)
CBR – A modified example

- **Initial retrieval**
  - A standard SQL query against a DB uses the following attributes:
    - Date of sale (within 24 months)
    - Distance (within 10 miles)
    - Living area (+ / - 25%)
    - Lot size (+ / - 50%)
    - Number of bedrooms (+/- 3)
    - Number of bathrooms (+/- 3)
**CBR – A modified example**

- **Similarity measure**
  - Weighted sum of attributes
  - Retrieved cases are ranked

---

<table>
<thead>
<tr>
<th>attribute</th>
<th>new case</th>
<th>retrieved case</th>
<th>comparison</th>
<th>weight</th>
<th>weighted sum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Months</td>
<td>X</td>
<td>6 months</td>
<td>75%</td>
<td>0.222</td>
<td>0.1665</td>
</tr>
<tr>
<td>Distance</td>
<td>X</td>
<td>0.2 miles</td>
<td>80%</td>
<td>0.222</td>
<td>0.1776</td>
</tr>
<tr>
<td>Area</td>
<td>2000</td>
<td>1800</td>
<td>90%</td>
<td>0.333</td>
<td>0.2997</td>
</tr>
<tr>
<td>Lot size</td>
<td>20000</td>
<td>35000</td>
<td>75%</td>
<td>0.111</td>
<td>0.8325</td>
</tr>
<tr>
<td>#Bedrooms</td>
<td>3</td>
<td>3</td>
<td>100%</td>
<td>0.056</td>
<td>0.056</td>
</tr>
<tr>
<td>#Bathrooms</td>
<td>2.5</td>
<td>2</td>
<td>80%</td>
<td>0.056</td>
<td>0.0448</td>
</tr>
</tbody>
</table>

**Similarity Measure** (Sum of Weighted Sum/Sum of Weights) = 0.8279

---

**Initial retrieval**

- Location
- Date of Sale
- Living Area
- Lot Size
- # Bedrooms
- # Bathrooms

**Compute similarity measure**
CBR – A modified example

- Adaptation rules
  - Adjust sales price to better reflect property value
  - Additional features cause the difference between subject and retrieved case
    - Fireplaces (subject - retrieved) * 2000
    - Quality (.02*sale price) for each level of difference: (Luxury > Excellent > Good > Average > Fair > Poor)
    - Lot Area (subject - retrieved) * 1
    -...
  - Rules obtained from engineering sessions with expert appraisers

Initial retrieval

Compute similarity measure

Apply adaptation rules

Final selection

Aggregate selected cases

Estimated value

Reliability

Justification
CBR – A modified example

- Aggregate selected cases
  - combined to produce an estimate of the value of the subject

<table>
<thead>
<tr>
<th>Retrieved cases</th>
<th>Adjusted price</th>
<th>Score</th>
<th>Weighted price</th>
</tr>
</thead>
<tbody>
<tr>
<td>113-012</td>
<td>197000</td>
<td>0.95</td>
<td>187150</td>
</tr>
<tr>
<td>306-008</td>
<td>202000</td>
<td>0.88</td>
<td>177760</td>
</tr>
<tr>
<td>093-011</td>
<td>196500</td>
<td>0.78</td>
<td>153270</td>
</tr>
<tr>
<td>685-046</td>
<td>192000</td>
<td>0.64</td>
<td>122880</td>
</tr>
<tr>
<td>847-984</td>
<td>201000</td>
<td>0.58</td>
<td>116580</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>3.83</td>
<td>757640</td>
</tr>
</tbody>
</table>

Final estimate = 757640/3.83 = 199900
Case Based Reasoning

CBR Applications

- **legal reasoning** (examples Hypo, JUDGE)
  - decision making in courts are based on legal precedents
- **diagnosis** (CASEY, Protos)
  - depends heavily on case histories and the doctor’s experience with other patients and their treatments
- **design** (Clavier)
  - successfully executed artifacts for a new situation
- **scheduling** (CABINS)
- **help-desk support** (Cascade, ReMind)
- **planning** (Chef)