Competence Guided Casebase Maintenance for Compositional Adaptation Applications

Ditty Mathew
Sutanu Chakraborti

Indian Institute of Technology, Madras

ICCBR 2016
Casebase Maintenance

Goal: Maintain a compressed casebase that can solve new problems effectively
Goal: Maintain a compressed casebase that can solve new problems effectively
Competence Guided Casebase Maintenance

- Competence of a CBR system is the range of target problems that the given system can solve
- Competence guided casebase maintenance system retains a case in the casebase if it is useful to solve many problems
- Thus it ensures that the casebase is highly competent in the global sense

*Smyth et al. Footprint-based Retrieval. In Case Based Reasoning Research and Development 1999
Competence Guided Casebase Maintenance

- Competence of a CBR system is the range of target problems that the given system can solve
- Competence guided casebase maintenance system retains a case in the casebase if it is useful to solve many problems
- Thus it ensures that the casebase is highly competent in the global sense
- Footprint-based approach* estimates a competent subset of the casebase

---

*Smyth et al. Footprint-based Retrieval. In Case Based Reasoning Research and Development 1999
Competence of a CBR system is the range of target problems that the given system can solve.

Competence guided casebase maintenance system retains a case in the casebase if it is useful to solve many problems.

Thus it ensures that the casebase is highly competent in the global sense.

Footprint-based approach* estimates a competent subset of the casebase.

However, Footprint based approach covers only the situation where a single case is adapted to solve a problem.

---

*Smyth et al. Footprint-based Retrieval. *In Case Based Reasoning Research and Development* 1999
Motivation

Competence Guided Casebase Maintenance

Single Case Adaptation
&
Compositional Adaptation
Footprint-based Approach

Uses Case Competence Model

Solves \( (c, t) \) \( \iff \) \( c \) can be retrieved and adapted for \( t \)

Coverage \( (c) = \{ c', \in C : \text{Solves}(c, c') \} \)

Reachability \( (c) = \{ c', \in C : \text{Solves}(c', c) \} \)

Coverage \( (c_3) = \{ c_2, c_4 \} \)

Reachability \( (c_2) = \{ c_1, c_3 \} \)

Reachability \( (c_4) = \{ c_3, c_5, c_6 \} \)
Footprint-based Approach

Uses Case Competence Model
Footprint-based Approach

Uses **Case Competence Model**

- \( \text{Solves}(c, t) \iff c \text{ can be retrieved and adapted for } t \)
Footprint-based Approach

Uses **Case Competence Model**

- Solves\((c, t) \iff c\) can be retrieved and adapted for \(t\)
- Coverage\((c) = \{c' \in C : \text{Solves}(c, c')\}\)
- Reachability\((c) = \{c' \in C : \text{Solves}(c', c)\}\)
Footprint-based Approach

Uses **Case Competence Model**

- \( \text{Solves}(c, t) \iff c \text{ can be retrieved and adapted for } t \)
- \( \text{Coverage}(c) = \{c' \in C : \text{Solves}(c, c')\} \)
- \( \text{Reachability}(c) = \{c' \in C : \text{Solves}(c', c)\} \)

Coverage(\(c_3\)) = \{c_2, c_4\}
Reachability(\(c_2\)) = \{c_1, c_3\}
Reachability(\(c_4\)) = \{c_3, c_5, c_6\}
Footprint-based Approach

Uses **Case Competence Model**

\[
\text{RelativeCoverage}(a) = \sum_{b \in \text{Coverage}(a)} \frac{1}{|\text{Reachability}(b)|}
\]

Coverage(c3) = \{c2, c4\}
Reachability(c2) = \{c1, c3\}
Reachability(c4) = \{c3, c5, c6\}
RelativeCoverage(c3) = \frac{1}{2} + \frac{1}{3}
Footprint-based Approach

Estimation of compact competent subset called *footprint set*

Table:

<table>
<thead>
<tr>
<th>Cases</th>
<th>Relative Coverage</th>
</tr>
</thead>
<tbody>
<tr>
<td>c1</td>
<td>2.25</td>
</tr>
<tr>
<td>c2</td>
<td>1.75</td>
</tr>
<tr>
<td>c4</td>
<td>0.5</td>
</tr>
<tr>
<td>c3</td>
<td>0.25</td>
</tr>
<tr>
<td>c5</td>
<td>0.25</td>
</tr>
</tbody>
</table>
Estimation of compact competent subset called *footprint set*

<table>
<thead>
<tr>
<th>Cases</th>
<th>Relative Coverage</th>
</tr>
</thead>
<tbody>
<tr>
<td>c1</td>
<td>2.25</td>
</tr>
<tr>
<td>c2</td>
<td>1.75</td>
</tr>
<tr>
<td>c4</td>
<td>0.5</td>
</tr>
<tr>
<td>c3</td>
<td>0.25</td>
</tr>
<tr>
<td>c5</td>
<td>0.25</td>
</tr>
</tbody>
</table>

Footprint set = $\emptyset$
Footprint-based Approach

Estimation of compact competent subset called *footprint set*

<table>
<thead>
<tr>
<th>Cases</th>
<th>Relative Coverage</th>
</tr>
</thead>
<tbody>
<tr>
<td>c1</td>
<td>2.25</td>
</tr>
<tr>
<td>c2</td>
<td>1.75</td>
</tr>
<tr>
<td>c4</td>
<td>0.5</td>
</tr>
<tr>
<td>c3</td>
<td>0.25</td>
</tr>
<tr>
<td>c5</td>
<td>0.25</td>
</tr>
</tbody>
</table>

Footprint set = ∅
Footprint-based Approach

Estimation of compact competent subset called *footprint set*

Footprint set = \{c1\}

<table>
<thead>
<tr>
<th>Cases</th>
<th>Relative Coverage</th>
</tr>
</thead>
<tbody>
<tr>
<td>c1</td>
<td>2.25</td>
</tr>
<tr>
<td>c2</td>
<td>1.75</td>
</tr>
<tr>
<td>c4</td>
<td>0.5</td>
</tr>
<tr>
<td>c3</td>
<td>0.25</td>
</tr>
<tr>
<td>c5</td>
<td>0.25</td>
</tr>
</tbody>
</table>
Footprint-based Approach

Estimation of compact competent subset called *footprint set*

Footprint set = \{c1\}

<table>
<thead>
<tr>
<th>Cases</th>
<th>Relative Coverage</th>
</tr>
</thead>
<tbody>
<tr>
<td>c1</td>
<td>2.25</td>
</tr>
<tr>
<td>c2</td>
<td>1.75</td>
</tr>
<tr>
<td>c4</td>
<td>0.5</td>
</tr>
<tr>
<td>c3</td>
<td>0.25</td>
</tr>
<tr>
<td>c5</td>
<td>0.25</td>
</tr>
</tbody>
</table>
Footprint-based Approach

Estimation of compact competent subset called *footprint set*

<table>
<thead>
<tr>
<th>Cases</th>
<th>Relative Coverage</th>
</tr>
</thead>
<tbody>
<tr>
<td>c1</td>
<td>2.25</td>
</tr>
<tr>
<td>c2</td>
<td>1.75</td>
</tr>
<tr>
<td>c4</td>
<td>0.5</td>
</tr>
<tr>
<td>c3</td>
<td>0.25</td>
</tr>
<tr>
<td>c5</td>
<td>0.25</td>
</tr>
</tbody>
</table>

Footprint set = \{c1\}
Footprint-based Approach

Estimation of compact competent subset called *footprint set*

<table>
<thead>
<tr>
<th>Cases</th>
<th>Relative Coverage</th>
</tr>
</thead>
<tbody>
<tr>
<td>c1</td>
<td>2.25</td>
</tr>
<tr>
<td>c2</td>
<td>1.75</td>
</tr>
<tr>
<td>c4</td>
<td>0.5</td>
</tr>
<tr>
<td>c3</td>
<td>0.25</td>
</tr>
<tr>
<td>c5</td>
<td>0.25</td>
</tr>
</tbody>
</table>

Footprint set = \{c1\}
Footprint-based Approach

In Compositional Adaptation applications

Footprint set = \{c1\}
Footprint-based Approach - Limitations

- Covers only single case adaptation
- Transitive coverage is not considered
Proposed a case competence model which covers *compositional adaptation* process (of which the single case adaptation is a special case)
We proposed a measure called retention score which quantifies the retention quality of a case in the casebase.
We proposed a measure called retention score which quantifies the retention quality of a case in the casebase.

- **CoveredCases(c)** include all cases that c solves either on its own, or in conjunction with other cases.
  
  - Eg: \(\text{CoveredCases}(c1) = \{c3, c4\}\)

- **SupportCases(c_i, c_j)** is the set of cases that are required to solve \(c_j\) using \(c_i\).
  
  - Eg: \(\text{SupportCases}(c1,c3) = \{c2\}\)
We proposed a measure called retention score which quantifies the retention quality of a case in the casebase.

**CoveredCases**(*c*) include all cases that *c* solves either on its own, or in conjunction with other cases.

- **Eg:** CoveredCases(*c*1) = {c3, c4}

**SupportCases**(*c*i, c*j*) is the set of cases that are required to solve *c*j using *c*i.

- **Eg:** SupportCases(*c*1, c3) = {c2}
RetentionScore Intuition

- A case has high retention score if it has
  - many covered cases
  - less number of support cases
RetentionPolicy Intuition (recursive formulation)

- A case has high retention score if it has
  - many covered cases with high retention score
  - less number of support cases with low retention score
For first iteration

\[
\text{RetentionScore}_0(c) = \sum_{c_i \in \text{CoveredCases}(c)} \frac{1/(1 + \text{No of alternate solutions that do not contain } c)}{1 + |\text{SupportCases}(c, c_i)|}
\]
For first iteration

RetentionScore_0(c) = \sum_{c_i \in \text{CoveredCases}(c)} \frac{1/(1 + \text{No of alternate solutions that do not contain c})}{1 + |\text{SupportCases}(c, c_i)|}

For estimating RetentionScore_0(c1)
For first iteration

\[
\text{RetentionScore}_0(c) = \sum_{c_i \in \text{CoveredCases}(c)} \frac{1}{1 + \text{No of alternate solutions that do not contain } c} \frac{1}{1 + |\text{SupportCases}(c, c_i)|}
\]

For estimating RetentionScore_0(c1)

- \text{CoveredCases}(c1) = \{c_2, c_3, c_4, c_5\}
For first iteration

\[
\text{RetentionScore}_0(c) = \sum_{c_i \in \text{CoveredCases}(c)} \frac{1}{1 + \text{No of alternate solutions that do not contain } c} \times \frac{1}{1 + |\text{SupportCases}(c, c_i)|}
\]

For estimating RetentionScore$_0(c_1)$
- \(\text{CoveredCases}(c_1) = \{c_2, c_3, c_4, c_5\}\)
- For a covered case \(c_4\),
Proposed Case Competence Model

For first iteration

\[
\text{RetentionScore}_0(c) = \sum_{c_i \in \text{CoveredCases}(c)} \frac{1}{1 + \text{No of alternate solutions that do not contain } c} \frac{1}{1 + |\text{SupportCases}(c, c_i)|}
\]

For estimating \(\text{RetentionScore}_0(c_1)\)
- \(\text{CoveredCases}(c_1) = \{c_2, c_3, c_4, c_5\}\)
- For a covered case \(c_4\),
  - \((c_1, c_2, c_5)\) form a solution
For first iteration

\[
\text{RetentionScore}_0(c) = \sum_{c_i \in \text{CoveredCases}(c)} \frac{1}{1 + \text{No of alternate solutions that do not contain } c} \frac{1}{1 + |\text{SupportCases}(c, c_i)|}
\]

For estimating RetentionScore$_0(c_1)$
- \( \text{CoveredCases}(c_1) = \{c_2, c_3, c_4, c_5\} \)
- For a covered case \( c_4 \),
  - \( (c_1, c_2, c_5) \) form a solution
  - \( (c_3) \) forms another solution
For first iteration

\[
\text{RetentionScore}_0(c) = \sum_{c_i \in \text{CoveredCases}(c)} \frac{1}{1 + \text{No of alternate solutions that do not contain } c} \frac{1}{1 + |\text{SupportCases}(c,c_i)|}
\]

For estimating RetentionScore\(_0(c_1)\)
- \(\text{CoveredCases}(c_1) = \{c_2, c_3, c_4, c_5\}\)
- For a covered case \(c_4\),
  - \((c_1, c_2, c_5)\) form a solution
  - \((c_3)\) forms another solution
  - \(\text{SupportCases}(c_1,c_4) = \{c_2,c_5\}\)
RetentionScore_0(c) = \sum_{c_i \in \text{CoveredCases}(c)} \frac{1}{1 + \text{No of alternate solutions that do not contain } c} \cdot \frac{1}{1 + |\text{SupportCases}(c, c_i)|}

RetentionScore_{k+1}(c) = \sum_{c_i \in \text{CoveredCases}(c)} \frac{\text{RetentionScore}_k(c_i)}{\sum_{c_j \in \text{SupportCases}(c, c_i)} (\text{RetentionScore}_k(c_j)) + 1}
Footprint$_{CA}$ Algorithm

- Modified Smyth’s footprint algorithm* to obtain footprint$_{CA}$ set
- Modified algorithm uses retention score instead of relative coverage

* Smyth et al. Footprint-based Retrieval. *In Case Based Reasoning Research and Development* 1999
Footprint$_{CA}$ Algorithm

<table>
<thead>
<tr>
<th>Cases</th>
<th>Retention Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>c1</td>
<td>2</td>
</tr>
<tr>
<td>c2</td>
<td>1.75</td>
</tr>
<tr>
<td>c3</td>
<td>1.29</td>
</tr>
<tr>
<td>c4</td>
<td>1.23</td>
</tr>
<tr>
<td>c5</td>
<td>1</td>
</tr>
</tbody>
</table>
Footprint$_{CA}$ Algorithm

\[ \text{Footprint}_{CA} \text{ set} = \emptyset \]

<table>
<thead>
<tr>
<th>Cases</th>
<th>Retention Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>c1</td>
<td>2</td>
</tr>
<tr>
<td>c2</td>
<td>1.75</td>
</tr>
<tr>
<td>c3</td>
<td>1.29</td>
</tr>
<tr>
<td>c4</td>
<td>1.23</td>
</tr>
<tr>
<td>c5</td>
<td>1</td>
</tr>
</tbody>
</table>
Footprint$_{CA}$ Algorithm

<table>
<thead>
<tr>
<th>Cases</th>
<th>Retention Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>c1</td>
<td>2</td>
</tr>
<tr>
<td>c2</td>
<td>1.75</td>
</tr>
<tr>
<td>c3</td>
<td>1.29</td>
</tr>
<tr>
<td>c4</td>
<td>1.23</td>
</tr>
<tr>
<td>c5</td>
<td>1</td>
</tr>
</tbody>
</table>

Footprint$_{CA}$ set = ∅
Footprint_{CA} Algorithm

Footprint_{CA} set = \{c1\}

<table>
<thead>
<tr>
<th>Cases</th>
<th>Retention Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>c1</td>
<td>2</td>
</tr>
<tr>
<td>c2</td>
<td>1.75</td>
</tr>
<tr>
<td>c3</td>
<td>1.29</td>
</tr>
<tr>
<td>c4</td>
<td>1.23</td>
</tr>
<tr>
<td>c5</td>
<td>1</td>
</tr>
</tbody>
</table>
Footprint$_{CA} \text{ Algorithm}$

Footprint$_{CA}$ set = \{c1\}

<table>
<thead>
<tr>
<th>Cases</th>
<th>Retention Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>c1</td>
<td>2</td>
</tr>
<tr>
<td>c2</td>
<td>1.75</td>
</tr>
<tr>
<td>c3</td>
<td>1.29</td>
</tr>
<tr>
<td>c4</td>
<td>1.23</td>
</tr>
<tr>
<td>c5</td>
<td>1</td>
</tr>
</tbody>
</table>
Footprint\textsubscript{CA} Algorithm

\begin{center}
\begin{tikzpicture}
\node[circle,fill=cyan,draw] (c1) at (0,0) {C1};
\node[circle,fill=cyan,draw] (c2) at (1,0) {C2};
\node[circle,fill=cyan,draw] (c3) at (1,1) {C3};
\node[circle,fill=cyan,draw] (c4) at (0,1) {C4};
\node[circle,fill=cyan,draw] (c5) at (-1,0) {C5};
\draw[->] (c1) to (c2);
\draw[->] (c1) to (c3);
\draw[->] (c1) to (c4);
\draw[->] (c1) to (c5);
\draw[->] (c2) to (c4);
\draw[->] (c3) to (c5);
\draw[->] (c4) to (c3);
\end{tikzpicture}
\end{center}

\begin{center}
\begin{tabular}{|l|l|}
\hline
Cases & Retention Score \\
\hline
\textcolor{red}{c1} & \textcolor{red}{2} \\
\hline
\textcolor{blue}{c2} & \textcolor{blue}{1.75} \\
\hline
\textcolor{blue}{c3} & \textcolor{blue}{1.29} \\
\hline
\textcolor{blue}{c4} & \textcolor{blue}{1.23} \\
\hline
\textcolor{blue}{c5} & \textcolor{blue}{1} \\
\hline
\end{tabular}
\end{center}

Footprint\textsubscript{CA} set = \{c1\}
Footprint$_{CA}$ Algorithm

Footprint$_{CA}$ set = \{c1, c3\}

<table>
<thead>
<tr>
<th>Cases</th>
<th>Retention Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>c1</td>
<td>2</td>
</tr>
<tr>
<td>c2</td>
<td>1.75</td>
</tr>
<tr>
<td>c3</td>
<td>1.29</td>
</tr>
<tr>
<td>c4</td>
<td>1.23</td>
</tr>
<tr>
<td>c5</td>
<td>1</td>
</tr>
</tbody>
</table>
Footprint_{CA} Algorithm

\[
\text{Footprint}_{CA} \text{ set } = \{ c_1, c_3 \}
\]

<table>
<thead>
<tr>
<th>Cases</th>
<th>Retention Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>c1</td>
<td>2</td>
</tr>
<tr>
<td>c2</td>
<td>1.75</td>
</tr>
<tr>
<td>c3</td>
<td>1.29</td>
</tr>
<tr>
<td>c4</td>
<td>1.23</td>
</tr>
<tr>
<td>c5</td>
<td>1</td>
</tr>
</tbody>
</table>
Footprint$_{CA}$ Algorithm

Footprint$_{CA}$ set = \{c1, c3\}

<table>
<thead>
<tr>
<th>Cases</th>
<th>Retention Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>c1</td>
<td>2</td>
</tr>
<tr>
<td>c2</td>
<td>1.75</td>
</tr>
<tr>
<td>c3</td>
<td>1.29</td>
</tr>
<tr>
<td>c4</td>
<td>1.23</td>
</tr>
<tr>
<td>c5</td>
<td>1</td>
</tr>
</tbody>
</table>
Footprint\textsubscript{CA} Algorithm

\begin{itemize}
\item Footprint\textsubscript{CA} set = \{c1, c3\}
\end{itemize}

\begin{table}
\centering
\begin{tabular}{|c|c|}
\hline
Cases & Retention Score \\
\hline
\textcolor{red}{c1} & 2 \\
\textcolor{red}{c2} & 1.75 \\
\textcolor{red}{c3} & 1.29 \\
\textcolor{red}{c4} & 1.23 \\
\textcolor{red}{c5} & 1 \\
\hline
\end{tabular}
\end{table}
Experiments

Synthetic Datasets

1. \[ y = x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 + x_8 + x_9 + x_{10} + \text{noise} \]
2. \[ y = x_1^4 + x_2^3 + x_3^2 + x_4 + \cos^2(x_5) + \text{noise} \]
3. \[ y = \sin(x_1 x_2) + \sqrt{x_3 x_4} + \cos^2(x_5) + x_6 x_7 + x_8 + x_9 + x_{10} + \text{noise} \]
Experiments

Synthetic Datasets

1. \[ y = x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 + x_8 + x_9 + x_{10} + \text{noise} \]
2. \[ y = x_1^4 + x_2^3 + x_3^2 + x_4 + \cos^2(x_5) + \text{noise} \]
3. \[ y = \sin(x_1 x_2) + \sqrt{x_3 x_4} + \cos^2(x_5) + x_6 x_7 + x_8 + x_9 + x_{10} + \text{noise} \]

- Each data instance is a case
- Each case is assumed to be solved by the combined solution of its K-nearest cases
Experiments

Synthetic Datasets

1. \[ y = x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 + x_8 + x_9 + x_{10} + \text{noise} \]
2. \[ y = x_1^4 + x_2^3 + x_3^2 + x_4 + \cos^2(x_5) + \text{noise} \]
3. \[ y = \sin(x_1x_2) + \sqrt{x_3x_4} + \cos^2(x_5) + x_6x_7 + x_8 + x_9 + x_{10} + \text{noise} \]

- Each data instance is a case
- Each case is assumed to be solved by the combined solution of its K-nearest cases
Evaluation - Footprint Size Analysis

Synthetic dataset 1

Footprint Size

Footprint\textsubscript{OR}

Footprint\textsubscript{CA}

$\text{k}$

$\text{Footprint Size}$

$\text{Synthetic dataset 1}$

Ditty Mathew, Sutanu Chakraborti

ICCBR 2016
**Evaluation - Casebase Coverage Analysis**

Casebase Coverage \((fp)\) = \[
\frac{\text{Cases that are solved by } fp}{\text{Casebase Size}}
\]

<table>
<thead>
<tr>
<th>Casebase Size</th>
<th>Synthetic data 1</th>
<th>Synthetic data 2</th>
<th>Synthetic data 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Synthetic data 1</th>
<th>Synthetic data 2</th>
<th>Synthetic data 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>% of Coverage</td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td></td>
<td></td>
</tr>
<tr>
<td>80</td>
<td></td>
<td></td>
</tr>
<tr>
<td>60</td>
<td></td>
<td></td>
</tr>
<tr>
<td>40</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Figure:** Casebase Coverage by Footprint \(OR\)
Sanity rate = \( \frac{|\text{footprint cases} \cap \text{kernel cases}|}{|\text{kernel cases}|} \times 100 \)

where,

- kernel cases* are obtained by repeatedly removing cases that do not solve any other cases until there are no such cases
- kernel cases cover the entire casebase

---

Synthetic Data 1

![Graph 1: Synthetic Data 1 - 1nn](image1)

![Graph 2: Synthetic Data 1 - 2nn](image2)

![Graph 3: Synthetic Data 1 - 4nn](image3)
Synthetic Data 2

Synthetic Data 2 - 1nn

Synthetic Data 2 - 2nn

Synthetic Data 2 - 4nn

Footprint<sub>CA</sub>  Footprint<sub>OR</sub>
Evaluation - Sanity Check

Synthetic Data 3

![Graph 1: Synthetic Data 3 - 1nn](image1)

![Graph 2: Synthetic Data 3 - 2nn](image2)

![Graph 3: Synthetic Data 3 - 4nn](image3)
- Encyclopedic resources like Wikipedia have less pedagogic value
- Concepts in Wikipedia (articles) are not arranged in a learning order
- An ideal textbook explains a concept before referring it which results in a sequential order for learning
- Sequencing concepts in Wikipedia may help online learners to fulfill their learning goal
Footprint\textsubscript{CA} in Tutoring Application

For Wikipedia

Atom

From Wikipedia, the free encyclopedia
(Redirected from Atoms)

For other uses, see Atom (disambiguation).

An atom is the smallest constituent unit of ordinary matter that has the properties of a chemical element.\textsuperscript{[3]} Every solid, liquid, gas, and plasma is composed of neutral or ionized atoms. Atoms are very small; typical sizes are around 100 pm (a ten-billionth of a meter, in the short scale).\textsuperscript{[7]} However, atoms do not have well-defined boundaries, and there are different ways to define their size that give different but close values.

Atoms are small enough that attempting to predict their behavior using classical physics - as if they were billiard balls, for example - gives noticeably incorrect predictions due to quantum effects. Through the development of physics, atomic models have incorporated quantum principles to better explain and predict the behavior.
Footprint$_{CA}$ in Tutoring Application

For Wikipedia

Atom

From Wikipedia, the free encyclopedia
(Redirected from Atoms)

For other uses, see Atom (disambiguation).

An atom is the smallest constituent unit of ordinary matter that has the properties of a chemical element. Every solid, liquid, gas, and plasma is composed of neutral or ionized atoms. Atoms are very small; typical sizes are around 100 pm (a ten-billionth of a meter, in the short scale) However, atoms do not have well-defined boundaries, and there are different ways to define their size that give different but close values.

Atoms are small enough that attempting to predict their behavior using classical physics - as if they were billiard balls, for example - gives noticeably incorrect predictions due to quantum effects. Through the development of physics, atomic models have incorporated quantum principles to better explain and predict the behavior.

Cases - Wikipedia articles
Footprint$_A$ in Tutoring Application

For Wikipedia

Atom

From Wikipedia, the free encyclopedia
(Redirected from Atoms)

For other uses, see Atom (disambiguation).

An atom is the smallest constituent unit of ordinary matter that has the properties of a chemical element.[3] Every solid, liquid, gas, and plasma is composed of neutral or ionized atoms. Atoms are very small; typical sizes are around 100 pm (a ten-billionth of a meter, in the short scale) [2] However, atoms do not have well-defined boundaries, and there are different ways to define their size that give different but close values.

Atoms are small enough that attempting to predict their behavior using classical physics - as if they were billiard balls, for example - gives noticeably incorrect predictions due to quantum effects. Through the development of physics, atomic models have incorporated quantum principles to better explain and predict the behavior.

Cases - Wikipedia articles

Problem Solution Pairs - (Article title A, Definition of article A)
Footprint$_{CA}$ in Tutoring Application

For Wikipedia

Atom

From Wikipedia, the free encyclopedia

(Redirected from Atoms)

For other uses, see Atom (disambiguation).

An atom is the smallest constituent unit of ordinary matter that has the properties of a chemical element. Every solid, liquid, gas, and plasma is composed of neutral or ionized atoms. Atoms are very small; typical sizes are around 100 pm (a ten-billionth of a meter, in the short scale) However, atoms do not have well-defined boundaries, and there are different ways to define their size that give different but close values.

Atoms are small enough that attempting to predict their behavior using classical physics - as if they were billiard balls, for example - gives noticeably incorrect predictions due to quantum effects. Through the development of physics, atomic models have incorporated quantum principles to better explain and predict the behavior.

Cases - Wikipedia articles

Problem Solution Pairs - (Article title A, Definition of article A)

We assume the first sentence of each article as its definition
For Wikipedia

Atom

From Wikipedia, the free encyclopedia
(Redirected from Atoms)

*For other uses, see Atom (disambiguation).*

An **atom** is the smallest constituent unit of ordinary **matter** that has the properties of a **chemical element**.[3] Every **solid**, **liquid**, **gas**, and **plasma** is composed of neutral or **ionized** atoms. Atoms are very small; typical sizes are around 100 pm (a ten-billionth of a meter, in the short scale).[2] However, atoms do not have well-defined boundaries, and there are different ways to **define their size** that give different but close values.

Atoms are **small enough** that attempting to predict their behavior using classical physics - as if they were billiard balls, for example - gives noticeably incorrect predictions due to **quantum effects**. Through the development of physics, atomic models have incorporated **quantum principles** to better explain and predict the behavior.
Footprint_{CA} in Tutoring Application

An example of casebase network from Wikipedia
Figure: Wikipedia Concept Network Example
Footprint\(_{CA}\) in Tutoring Application

**Figure:** Wikipedia Concept Network Example

Footprint\(_{CA}\) set - \{Atom, Chemical Element, Chemical Compound\}

<table>
<thead>
<tr>
<th>Concepts</th>
<th>Retention Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Atom</td>
<td>2.0</td>
</tr>
<tr>
<td>Matter</td>
<td>1.19</td>
</tr>
<tr>
<td>Chemical Element</td>
<td>1.18</td>
</tr>
<tr>
<td>Chemical Compound</td>
<td>1.12</td>
</tr>
<tr>
<td>Chemical Bond</td>
<td>1</td>
</tr>
</tbody>
</table>
Evaluation on Dictionary and Wikipedia Datasets

ldoce - Longman dictionary, wn - WordNet, wikiAI - Wikipedia (A. I. Category)
Evaluation on Dictionary and Wikipedia Datasets

\begin{align*}
\text{Sanity rate} &= \frac{|\text{footprint cases} \cap \text{kernel cases}|}{|\text{kernel cases}|} \times 100
\end{align*}

ldoce - Longman dictionary, wn - WordNet, wikiAI - Wikipedia (A.I. Category)
Conclusion

- Retention Score orders cases based on retention quality
- Modified footprint algorithm estimates competent compressed casebase using retention score ordering
- Experimented the idea in synthetic datasets and tutoring application


Thank You!!
Thank You!!
Questions??
Estimating Kernel Cases
Estimating Kernel Cases

C4 -> C3
C3 -> C2
C2 -> C1
C1 -> C4
Estimating Kernel Cases

C1 \rightarrow C3
C3 \rightarrow C2
C2 \rightarrow C1
Estimating Kernel Cases

C1

C2