Towards a Principle-based Approach for Case-based Reasoning

Leila Amgoud, Vivien Beuselinck

ANITI, Université Fédérale (France)

09/10/2022

Case-based reasoning (CBR) is the process of solving new problems based on the solutions of **similar** past problems.

- Many practical applications (Law, Medicine, ...).
- A plethora of CBR models.

- Few works on theoretical foundations of CBR.
- Lack of formal comparison with non monotonic reasoning (NMR).

Case-based reasoning (CBR) is the process of solving new problems based on the solutions of **similar** past problems.

- Many practical applications (Law, Medicine, ...).
- A plethora of CBR models.

- Few works on theoretical foundations of CBR.
- Lack of formal comparison with non monotonic reasoning (NMR).

The paper bridges the two gaps.

Plan

- Introduction
- Basic assumption of CBR => 4 notions
- Axioms for CBR
- Summary

Example [1]

Cases	Years old	Power	Mileage	Equip	Shape	Price
C_1	1	1300	20 000	poor	good	8000
C_2	2	1600	30 000	excellent	poor	7000
C_3	2	1600	40 000	good	good	5000
C_4	3	1500	60 000	excellent	poor	5000
C_n	2	1600	50 000	poor	good	?

[1] Didier Dubois, Francesc Esteva, Pere Garcia, Lluís Godo, Ramón López de Mántaras, Henri Prade. Case-Based Reasoning: A Fuzzy Approach Fuzzy Logic in Artificial Intelligence. IJCAI Workshop 1997

Features and instances

- Features : $\mathcal{F} = \{f_1, \dots, f_n, \mathtt{f}\}$
- Domains : $dom(x)$ returns the domain of feature $x \in \mathcal{F}$.
  $dom(\mathtt{f})$ is the set of possible outcomes
- Literals : Every pair (f, v) such that $f \in \mathcal{F} \setminus \{\mathtt{f}\}$ and $v \in \mathrm{dom}(f)$
- Instances : A set of literals, where each feature $f_1, \dots, f_n$ appears exactly once.
  `Inst` : The set of all possible instances.
- Case : $c = \langle I, v \rangle$ such that $I \in \mathtt{Inst}$ and $v \in \mathrm{dom}(\mathtt{f}) \cup \{?\}$
  Past Case : When $v \in dom(f)$
  New Case : when $v = ?$
- Case Base : A set of n past cases.

Two similarity measures

- ($\mathbf{S}^i$) : For comparing attributes-values.
- ($\mathbf{S}^o$) : For comparing prices.

Two thresholds :

- x is dissimilar to y iff $\mathbf{S}^i(x, y) < \delta^i$
- v is dissimilar to v' iff $\mathbf{S}^o(v, v') < \delta^o$

Example :

$$\mathbf{S}^o(u, v) = \begin{cases} 1 & \text{if } |u - v| \leq 500 \\ 0 & \text{if } |u - v| \geq 2000 \\ 1 - \frac{1}{1500} * (|u - v| - 500) & \text{if } 500 < |u - v| < 2000 \end{cases}$$

Theory

A theory is a tuple $\mathbf{T} = \langle \mathcal{F}, \text{dom}, \mathbf{S}^i, \mathbf{S}^o, \delta^i, \delta^o \rangle$.

Basic assumption of CBR

"The more similar the cases, the more similar their outcomes".

Four notions that capture the assumption.

- Consistency
- Strong Coherence
- Weak Coherence
- Regularity

Idea: Similar input data receive similar outcomes

Consistency

A case base Σ is ***consistent*** *iff $\forall \langle I, v \rangle, \langle I', v' \rangle \in \Sigma$, if $\mathbf{S}^i(I, I') = 1$ then $\mathbf{S}^o(v, v') = 1$. It is inconsistent otherwise.*

Property

If a case base Σ is consistent, then $\forall \langle I, v \rangle, \langle I', v' \rangle \in \Sigma$, if $I = I'$, then $\mathbf{S}^o(v, v') = 1$.

Strong Coherence

*A case base Σ is **strongly coherent** iff $\forall \langle I, v\rangle, \langle I', v'\rangle \in \Sigma$, $\mathbf{S}^i(I, I') \leq \mathbf{S}^o(v, v')$.*

Example

Strong Coherence is violated because $S^o(C_1, C_3) = 0$ (Case $|u - v| > 2000$ with $|8000 - 5000| = 3000$)

Property

Let Σ be a strongly coherent case base. For all $\langle I, v\rangle, \langle I', v'\rangle \in \Sigma$, if $\mathbf{S}^i(I, I') = 1$, then $\mathbf{S}^o(v, v') = 1$.

Idea: similar input data should receive similar outcomes

Weak Coherence

*A case-base Σ is **weakly coherent** iff $\forall\langle I, v\rangle, \langle I', v'\rangle \in \Sigma$, if $\mathbf{S}^i(I, I') \geq \delta^i$, then $\mathbf{S}^o(v, v') \geq \delta^o$.*

Example

With $\delta_i = \frac{4}{5}$ and $\delta_o = \frac{3}{5}$, the case base satisfies Weak Coherence.

Suppose we have a case base on student grades. There are 4 attributes corresponding to courses with a mark; the outcome is an appreciation whose range consists of 4 qualitative levels. The similarity measure $\mathbf{S}^i$ takes the minimal value returned by $\mathbf{S}$ on the four courses. Assume Σ contains two students who got respectively $I = \langle 20, 20, 20, 20 \rangle$ and $v =$ "excellent" as global appreciation, and $I' = \langle 20, 20, 15, 15 \rangle$ with appreciation $v' =$ "good". Hence, $\mathbf{S}^i(I, I') = 0.75$ and $\mathbf{S}^o(v, v') = \frac{2}{3}$. Note that the base is not strongly coherent. In order to be coherent, $\mathbf{S}^o(v, v')$ should be equal to 1, which is not reasonable in the example as the two instances are different and deserve different appreciations. Furthermore, the scale of $\mathbf{S}^o$ does not have an intermediate value between $\frac{2}{3}$ and 1.

Regularity

*A case-base Σ is **regular** iff $\forall \langle I, v\rangle, \langle I', v'\rangle, \langle I'', v''\rangle \in \Sigma$, if $\mathbf{S}^i(I, I') \geq \mathbf{S}^i(I, I'')$ then $\mathbf{S}^o(v, v') \geq \mathbf{S}^o(v, v'')$.*

Example

Regularity is violated because $S^i(C_2, C_3) > S^i(C_2, C_1)$ but $S^o(C_2, C_3) < S^o(C_2, C_1)$.

Links

- Strong coherence implies consistency
- Strong coherence implies Weak Coherence if $\delta^i \geq \delta^o$
- Others are independent.

CBR model

Let $\mathbf{T} = \langle \mathcal{F}, \mathtt{dom}, \mathbf{S}^i, \mathbf{S}^o, \delta^i, \delta^o \rangle$ *be a theory. A CBR model is a function* R *mapping every case base* Σ *and new case* $\langle I, ? \rangle$ *into a set* $O \subseteq \mathtt{dom}(\mathtt{f}) \cup \{\mathtt{Und}\}$ *such that* $O \neq \emptyset$ *and either* $O = \{\mathtt{Und}\}$ *or* $O \subseteq \mathtt{dom}(\mathtt{f})$*. We write* $\Sigma \oplus \langle I, ? \rangle \mathrel{|\!\sim}_{\mathbf{T},\mathrm{R}} O$*.*

CBR AXIOMS

Strong Completeness

$\Sigma \oplus \langle I, ? \rangle \mathrel{|\!\sim}_{\mathbf{T},\mathrm{R}} O$ *with* $O \subseteq \mathtt{dom(f)}$.

Weak Completeness

$\Sigma \oplus \langle I, ? \rangle \mathrel{|\!\sim}_{\mathbf{T},\mathrm{R}} \{\mathrm{Und}\}$ *iff* $\forall \langle I, v \rangle \in \Sigma$, $\mathbf{S}^i(I_n, I) < \delta^i$.

Property

- *if a model* **R** *satisfies weak completeness, then* $\emptyset \oplus \langle I, ? \rangle \mathrel{|\!\sim}_{\mathbf{T},\mathrm{R}} \{\mathrm{Und}\}$.
- *Strong completeness and weak completeness are incompatible*

CBR AXIOMS (CONT.)

Consistency

Let Σ be consistent and $\Sigma \oplus \langle I, ?\rangle \mid\!\sim_{\mathbf{T},\mathrm{R}} O$ such that $O \neq \{\mathrm{Und}\}$. For any $v \in O$, $\Sigma \cup \{\langle I_n, v\rangle\}$ is consistent.

Strong Coherence

Let Σ be strongly coherent and $\Sigma \oplus \langle I, ?\rangle \mid\!\sim_{\mathbf{T},\mathrm{R}} O$ such that $O \neq \{\mathrm{Und}\}$. For any $v \in O$, $\Sigma \cup \{\langle I_n, v\rangle\}$ is strongly coherent.

Weak Coherence

Let Σ be weakly coherent and $\Sigma \oplus \langle I, ? \rangle \mid\!\sim_{\mathbf{T},\mathrm{R}} O$ such that $O \neq \{\mathrm{Und}\}$. For any $v \in O$, $\Sigma \cup \{\langle I_n, v \rangle\}$ is weakly coherent.

Property

Let $\mathbf{T} = \langle \mathcal{F}, \mathrm{dom}, \mathbf{S}^i, \mathbf{S}^o, \delta^i, \delta^o \rangle$ be a theory such that $\delta^i \geq \delta^o$. If a CBR model satisfies strong coherence, then it satisfies weak coherence.

Regularity

Let Σ be regular and $\Sigma \oplus \langle I, ? \rangle \mid\!\sim_{\mathbf{T},\mathrm{R}} O$ such that $O \neq \{\mathrm{Und}\}$. For any $v \in O$, $\Sigma \cup \{\langle I_n, v \rangle\}$ is regular.

Monotonicity

$$\left\{\begin{array}{l}\Sigma \oplus \langle I, ?\rangle \mid\!\sim_{\mathbf{T},\mathrm{R}} O \\ \\ \Sigma \subseteq \Sigma'\end{array}\right. \implies \Sigma' \oplus \langle I, ?\rangle \mid\!\sim_{\mathbf{T},\mathrm{R}} O$$

Property

If a CBR model satisfies weak completeness, then it violates monotonicity.

Cautious Monotonicity

$$\begin{cases} \Sigma \oplus \langle I, ?\rangle \mid\!\sim_{\mathbf{T},\mathrm{R}} \{v\} \\ \\ \Sigma \oplus \langle I', ?\rangle \mid\!\sim_{\mathbf{T},\mathrm{R}} \{v'\} \end{cases} \implies \Sigma \cup \{\langle I, v\rangle\} \oplus \langle I', ?\rangle \mid\!\sim_{\mathbf{T},\mathrm{R}} \{v'\}$$

Property

Strong coherence and cautious monotonicity are incompatible.

SUMMARY

Conclusion

- A proposal of some axioms for CBR.
- Highlighting some differences with non Monotonic Reasoning.

Challenges

- More axioms for CBR.
- Characterization of formal models that satisfy the axioms.