Representing Case Variations for Learning General and Specific Adaptation Rules

Fadi Badra, Jean Lieber

Orpailleur Group
LORIA (UMR 7503 CNRS–INPL–INRIA-Nancy 2–UHP)
BP 239, 54506 Vandœuvre-lès-Nancy, France
{badra,lieber}@loria.fr

Contributions of the paper

- A **representation of variations between cases** in case-based reasoning is proposed.
- The task of **adaptation knowledge acquisition** is formalized as a **problem of learning by generalization**.
- First experiments were run in the oncology domain.

Case-Based Reasoning

Solving a target problem using a set of already solved problems

- A new problem is called a target problem ($\mathtt{tgt}$).
- The case base is the set of the source cases.
- A source case is a pair

$$(\mathtt{srce}, \mathtt{Sol}(\mathtt{srce})) \text{ such that } \begin{cases} \mathtt{srce} \text{ is a source problem} \\ \mathtt{Sol}(\mathtt{srce}) \text{ is a solution of } \mathtt{srce} \end{cases}$$

Case-Based Reasoning

Example in oncology domain

- Problem-solving task: recommending a treatment to a patient given its description.
- The case base contains descriptions of medical situations.
- $\mathtt{srce}$ is a patient description.
- $\mathtt{Sol}(\mathtt{srce})$ is a treatment recommendation.
- $\mathtt{tgt}$ is the description of a new patient.

Case-Based Reasoning: Decomposition

Transformational
analogy

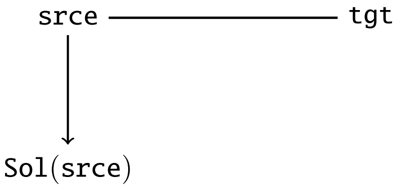
Case-Based Reasoning: Decomposition

tgt

Transformational
analogy

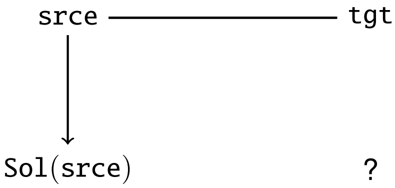
Case-Based Reasoning: Decomposition

Transformational analogy



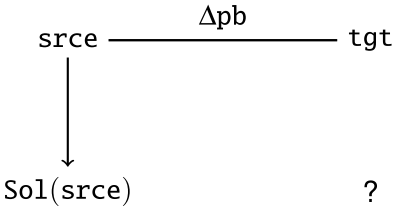
Case-Based Reasoning: Decomposition

Transformational analogy



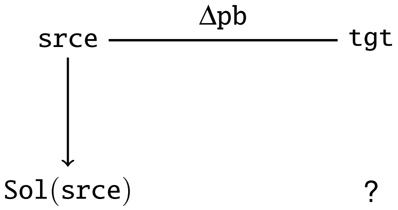
Case-Based Reasoning: Decomposition

Transformational analogy



Case-Based Reasoning: Decomposition

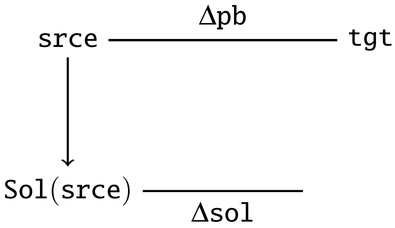
Transformational analogy



$$\mathsf{AK} : \Delta\mathtt{pb} \mapsto \Delta\mathtt{sol}$$

Case-Based Reasoning: Decomposition

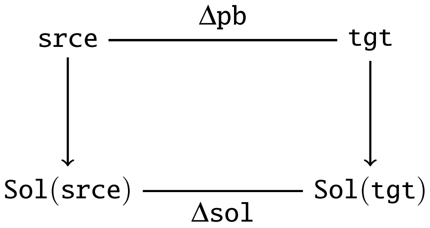
Transformational analogy



$\mathsf{AK} : \Delta\mathtt{pb} \mapsto \Delta\mathtt{sol}$

Case-Based Reasoning: Decomposition

Transformational analogy

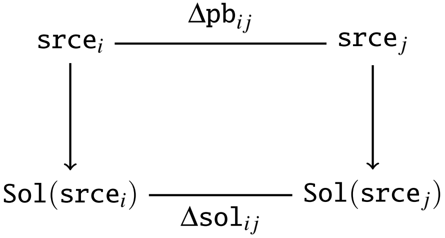


AK : $\Delta\mathtt{pb} \mapsto \Delta\mathtt{sol}$

Adaptation Knowledge Acquisition

Following K. Hanney's approach (1996)

Learn AK
from variations
between cases
in the case base

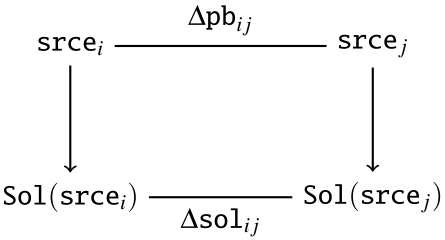


$$(\Delta\mathtt{pb}_{ij}, \Delta\mathtt{sol}_{ij})_{ij} \mapsto \text{AK}$$

Adaptation Knowledge Acquisition

Following K. Hanney's approach (1996)

Learn AK
from variations
between cases
in the case base



$$(\Delta\mathtt{pb}_{ij}, \Delta\mathtt{sol}_{ij})_{ij} \mapsto \text{AK}$$

Outline of the approach

1. Represent the variations between cases
2. Design an algorithm that learns AK by generalizing from these variations

The Generalization Problem

As stated by T. Mitchell (1990)

"Structure a set of individuals by generalizing beyond observed data".

Given

1. a set of instances
2. a language of instances
3. a language of generalizations
4. some matching predicates

Determine

a generalization consistent with the training instances.

The Generalization Problem

As stated by T. Mitchell (1990)

"Structure a set of individuals by generalizing beyond observed data".

Given

1. a set of instances
2. a language of instances
3. a language of generalizations
4. some matching predicates

the pairs of all distinct source cases of the case base

Determine

a generalization consistent with the training instances.

The Generalization Problem

As stated by T. Mitchell (1990)

"Structure a set of individuals by generalizing beyond observed data".

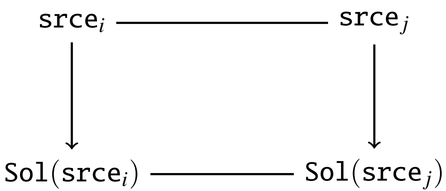
Given

1. a set of instances
2. a language of instances
3. a language of generalizations
4. some matching predicates

Determine

a generalization consistent with the training instances.

A Language of Instances



An instance:

- is a pair of distinct source cases
- is an element of $\mathscr{L}_{\mathtt{pb}} \times \mathscr{L}_{\mathtt{sol}} \times \mathscr{L}_{\mathtt{pb}} \times \mathscr{L}_{\mathtt{sol}}$

The Generalization Problem

As stated by T. Mitchell (1990)

"Structure a set of individuals by generalizing beyond observed data".

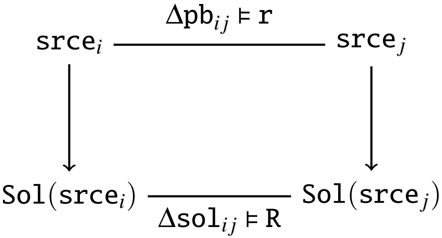
Given

1. a set of instances
2. a language of instances
3. a language of generalizations
4. some matching predicates

Determine

a generalization consistent with the training instances.

A Language of Generalizations



Representing Variations

- **Variations** Δpb_{ij} and Δsol_{ij} are represented as **binary relations**
- Generalizing over these variations is achieved by considering more general relations $r \in \mathscr{L}_{\Delta pb}$ and $R \in \mathscr{L}_{\Delta sol}$.

A Language of Generalizations

Example

To represent the variation $\Delta\texttt{pb}_{ij}$ between the source problems

$$\texttt{srce}_i = (\texttt{age}, 28) \wedge \ldots$$
$$\texttt{srce}_j = (\texttt{age}, 41) \wedge \ldots$$

we may define the relations $\texttt{age}^{\neq}$ or $\texttt{age}^{<}$ with $\begin{cases} \texttt{srce}_i\ \texttt{age}^{\neq}\ \texttt{srce}_j \\ \texttt{srce}_i\ \texttt{age}^{<}\ \texttt{srce}_j \end{cases}$

Choosing Relations

Example

$$\text{if} \quad \begin{array}{l} \mathtt{srce}_i = (\mathtt{age}, [16;45]) \wedge \ldots \\ \mathtt{srce}_j = (\mathtt{age}, [65;70]) \wedge \ldots \end{array}$$

$$\text{then} \quad \mathtt{srce}_i \; \mathtt{age}^b \; \mathtt{srce}_j$$

b is the Allen's relation $before$:

$$[a_1;b_1] \; b \; [a_2;b_2] \text{ iff } b_1 < a_2$$

- The choice of relations of $\mathscr{L}_{\Delta \mathrm{pb}}$ and $\mathscr{L}_{\Delta \mathtt{sol}}$ is a knowledge acquisition from experts issue.
- It constitutes a representational bias in the learning process.

Semantics of relations

- $\mathtt{r} \in \mathscr{L}_{\Delta \mathrm{pb}}$ is interpreted as a subset $\mathrm{Ext}(\mathtt{r})$ of $\mathscr{L}_{\mathrm{pb}} \times \mathscr{L}_{\mathrm{pb}}$.

Examples

$\mathtt{r}$	$\mathrm{Ext}(\mathtt{r})$
$\langle \mathtt{srce}_i, \mathtt{srce}_j \rangle$	$\{(\mathtt{srce}_i, \mathtt{srce}_j)\}$
$\mathtt{age}^{\neq}$	$(\mathtt{srce}_i, \mathtt{srce}_j)$ for which ages differ
$\mathtt{age}^{<}$	$(\mathtt{srce}_i, \mathtt{srce}_j)$ for which the age increases

- The semantics induce a generalization relation $\models$.

$$\mathtt{age}^{<} \models \mathtt{age}^{\neq} \text{ holds since } \mathrm{Ext}(\mathtt{age}^{<}) \subseteq \mathrm{Ext}(\mathtt{age}^{\neq})$$

The Generalization Problem

As stated by T. Mitchell (1990)

"Structure a set of individuals by generalizing beyond observed data".

Given

1. a set of instances
2. a language of instances
3. a language of generalizations
4. some matching predicates

$\mathtt{r} \in \mathscr{L}_{\Delta\mathrm{pb}}$ covers $(\mathtt{srce}_i, \mathtt{srce}_j)$
if $(\mathtt{srce}_i, \mathtt{srce}_j) \in \mathrm{Ext}(\mathtt{r})$

$\mathrm{R} \in \mathscr{L}_{\Delta\mathtt{sol}}$ covers $(\mathtt{Sol}(\mathtt{srce}_i), \mathtt{Sol}(\mathtt{srce}_j))$
if $(\mathtt{Sol}(\mathtt{srce}_i), \mathtt{Sol}(\mathtt{srce}_j)) \in \mathrm{Ext}(\mathrm{R})$

Determine

a generalization consistent with the training instances.

The Generalization Problem

As stated by T. Mitchell (1990)

"Structure a set of individuals by generalizing beyond observed data".

Given

1. a set of instances
2. a language of instances
3. a language of generalizations
4. some matching predicates

Determine

a generalization consistent with the training instances.

Generalizing over pairs of source cases

Definition

The variation Δpb_{ij} between two source problems $\texttt{srce}_i$ and $\texttt{srce}_j$ is represented by the relation $\Delta\text{pb}_{ij} = \langle \texttt{srce}_i, \texttt{srce}_j \rangle$.

- Any relation $\texttt{r}$ such that $\langle \texttt{srce}_i, \texttt{srce}_j \rangle \vDash \texttt{r}$ is a generalization of Δpb_{ij}.

Example

$$\text{if} \quad \begin{aligned} \texttt{srce}_i &= (\text{age}, 28) \wedge \ldots \\ \texttt{srce}_j &= (\text{age}, 41) \wedge \ldots \end{aligned}$$

$$\text{then} \quad \Delta\text{pb}_{ij} = \langle \texttt{srce}_i, \texttt{srce}_j \rangle \vDash \text{age}^{<}$$

Learning Generalizations

Learning Adaptation Knowledge

- Associate to each pair of source cases the set of relations that cover it, i.e.,

$$\{\mathtt{r} \in \mathscr{L}_{\Delta\mathrm{pb}} \mid \Delta\mathrm{pb}_{ij} = \langle \mathtt{srce}_i, \mathtt{srce}_j \rangle \vDash \mathtt{r}\}$$

$$\text{and} \quad \{\mathtt{R} \in \mathscr{L}_{\Delta\mathtt{sol}} \mid \Delta\mathtt{sol}_{ij} = \langle \mathtt{Sol}(\mathtt{srce}_i), \mathtt{Sol}(\mathtt{srce}_j) \rangle \vDash \mathtt{R}\}$$

- Extract most frequent sets of relations and interpret them as adaptation rules

Learning Generalizations

- Learning algorithm = frequent closed itemset extraction algorithm : CHARM (Zaki, 2002) implemented in the CORON platform (Szathmary, 2005) `http://coron.loria.fr`
- CHARM inputs a formal context $\mathscr{C}$ in which
  - rows represent pairs of source cases
  - columns represent relations

Learning Generalizations

Example (simplified)

Let

$$\begin{aligned} \texttt{srce}_i &= (\texttt{age}, [16; 45]) \wedge \ldots \\ \texttt{Sol}(\texttt{srce}_i) &= (\texttt{nb-of-FEC-cycles}, 10) \wedge (\texttt{dose-of-FEC}, 100) \wedge \ldots \end{aligned}$$

and

$$\begin{aligned} \texttt{srce}_j &= (\texttt{age}, [65; 70]) \wedge \ldots \\ \texttt{Sol}(\texttt{srce}_i) &= (\texttt{nb-of-FEC-cycles}, 5) \wedge (\texttt{dose-of-FEC}, 50) \wedge \ldots \end{aligned}$$

be two cases of the case base. Then $\mathscr{C}$ contains the row

	...	$\texttt{age}^b$	...	$\texttt{nb-of-FEC-cycles}^>$	...	$\texttt{dose-of-FEC}^>$	...
...							
o_{ij}		x		x		x	...
...							

Learning Generalizations

Example (continued)

With the following context $\mathscr{C}$:

	...	$\texttt{age}^b$	...	$\texttt{nb-of-FEC-cycles}^>$	...	$\texttt{dose-of-FEC}^>$	...
...							
o_{ij}		x		x		x	...
...							

the itemset

$$\mathscr{I} = \{\texttt{age}^b, \texttt{nb-of-FEC-cycles}^>, \texttt{dose-of-FEC}^>\}$$

generalizes a set of pairs of source cases among which is o_{ij}.

A frequent itemset:

- represents a set of relations between source cases
- is interpreted as an adaptation rule.

Adaptation Rules

Definition

An adaptation rule is an ordered pair $(\mathtt{r}, \mathtt{R}) \in \mathscr{L}_{\Delta\mathrm{pb}} \times \mathscr{L}_{\Delta\mathtt{sol}}$. It is interpreted as follows:

if $\langle \mathtt{srce}, \mathtt{tgt} \rangle \vDash \mathtt{r}$

then $\mathtt{Sol}(\mathtt{tgt})$ is such that $\langle \mathtt{Sol}(\mathtt{srce}), \mathtt{Sol}(\mathtt{tgt}) \rangle \vDash \mathtt{R}$

Results

Example of result

The itemset

$$\mathscr{I} = \{\texttt{age}^b, \texttt{nb-of-FEC-cycles}^>, \texttt{dose-of-FEC}^>\}$$

gives the general adaptation rule

$$\texttt{AR} = (\texttt{r}, \texttt{R}) \text{ where } \begin{cases} \texttt{r} = \texttt{age}^b \\ \texttt{R} = \texttt{nb-of-FEC-cycles}^> \wedge \texttt{dose-of-FEC}^> \end{cases}$$

Interpretation:

When the age of the patient increases, the number of cycles of chemotherapy decreases and the dose per cycle decreases.

Structuring the Result Set

Generalization relation between adaptation rules

$$(\mathtt{r}_1, \mathtt{R}_1) \vDash (\mathtt{r}_2, \mathtt{R}_2) \text{ iff } \mathtt{r}_1 \vDash \mathtt{r}_2 \text{ and } \mathtt{R}_1 \vDash \mathtt{R}_2$$

Example

$$\mathtt{AR}' = (\mathtt{r}', \mathtt{R}') \text{ where } \begin{cases} \mathtt{r}' = \mathtt{age}^{\neq} \\ \mathtt{R}' = \mathtt{nb\text{-}of\text{-}FEC\text{-}cycles}^{\neq} \wedge \mathtt{dose\text{-}of\text{-}FEC}^{\neq} \end{cases}$$

is more general than (and so more frequent than)

$$\mathtt{AR} = (\mathtt{r}, \mathtt{R}) \text{ where } \begin{cases} \mathtt{r} = \mathtt{age}^{b} \\ \mathtt{R} = \mathtt{nb\text{-}of\text{-}FEC\text{-}cycles}^{>} \wedge \mathtt{dose\text{-}of\text{-}FEC}^{>} \end{cases}$$

Structuring the Result Set

A Hierarchy of Adaptation Rules

- The learning algorithm generates a hierarchy of adaptation rules.
- The generality relation $\models$ **structures the result set** and allows to navigate in it.
- The most general rules are also the most frequent ones.

Conclusion

- A formalization of variations between cases in case-based reasoning using a language of binary relations between cases
- A learning algorithm that constructs a hierarchy of relations which can be used to determine and organize candidate adaptation rules
- First experiments run in the oncology domain

Ongoing Work

- More experiments in order to validate the approach
- Designing tools to navigate in the extracted adaptation rules
- Exploiting the hierarchy of relations in a given CBR session to determine which adaptation rules apply to a given target problem (classification procedure)
- Study the composition of adaptation rules