Argumentation-based learning from communication

IIIA-CSIC

Enric Plaza

A-MAIL framework: combines inductive learning with computational argumentation for multiagent systems (MAS):
- Uses **Inductive Learning**: to create arguments and counterargument
- Uses **Computational Argumentation** to refine and change inductive hypothesis
- Learning fom an **Oracle** (Teacher) is a particular case of MAS in which learner and teacher argue about learnt concepts

«Oracles» should Argue

CLAIM 1: Lifelong learning requires a more complex relation with "oracle" about right and wrong: **argumentation** to give a rihcer feedback and allow system autonomy (rejecting arguments on empirical individual basis)

CLAIM 2: Lifelong learning lives in social context: we need learning that aligns with external individuals and systems (ecology determines 'correctness'): argumentation to mutually exchange and adapt learnt concepts.

Computational Argumentation

Two types of arguments:

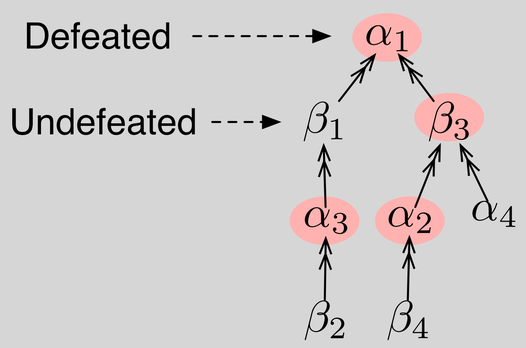
- Example Arguments: $\langle e, \overline{C} \rangle$
- Rule Arguments: $\langle h, \overline{C} \rangle$

$$\overline{C} \in \{+, -\}$$

Attack:

$$\langle h_1, \overline{C} \rangle \twoheadrightarrow \langle h_2, \widehat{C} \rangle \iff \overline{C} = \neg\widehat{C} \wedge h_2 \sqsubset h_1$$
$$\langle e, \overline{C} \rangle \twoheadrightarrow \langle h, \widehat{C} \rangle \iff \overline{C} = \neg\widehat{C} \wedge h \sqsubseteq e$$

Argumentation Trees:



Multiagent Inductive Learning (MAIL)

Given:

- A set of agents $\{A_1, ..., A_m\}$
- With their training sets $E_1, ..., E_m$
- A target concept $C : \mathcal{E} \rightarrow \{+, -\}$

Each agent wants to find a hypothesis:

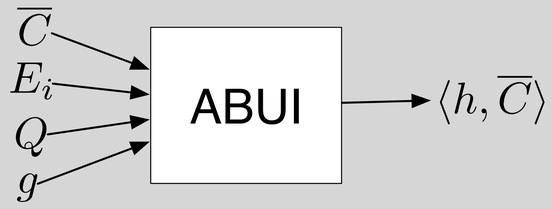
$$H_i \mid \forall_{e \in E_1 \cup ... \cup E_m} H_i(e) = C(e)$$

(Here we focus in 2-agent scenario)

Generating Arguments via Inductive Learning

ABUI: Argumentation-based Bottom-Up Inductive algorithm

Given:
- A target concept
- A set of examples
- A set of agreed upon arguments
- A generalization



ABUI generates a rule for the target concept which is not attacked by any agreed upon argument, more specific than the given generalization, and with confidence above a given threshold

Confidence: $B(\langle h, \overline{C} \rangle) = \dfrac{Y + 1}{Y + N + 2}$

Only arguments with confidence over a threshold allowed in A-MAIL

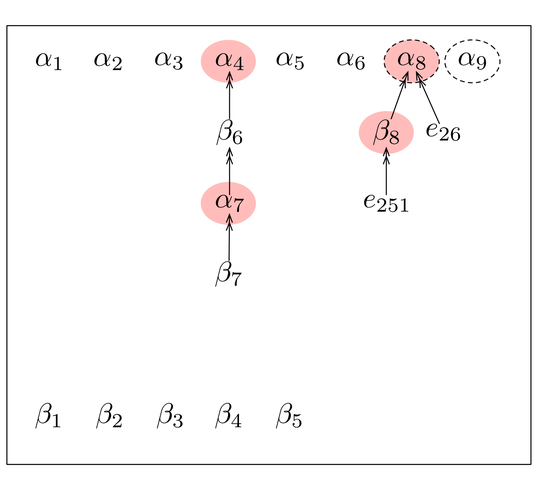
Example of an argumentation run

$$R_1^0 = \{\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5, \alpha_6\}$$
$$R_2^0 = \{\beta_1, \beta_2, \beta_3, \beta_5\}$$

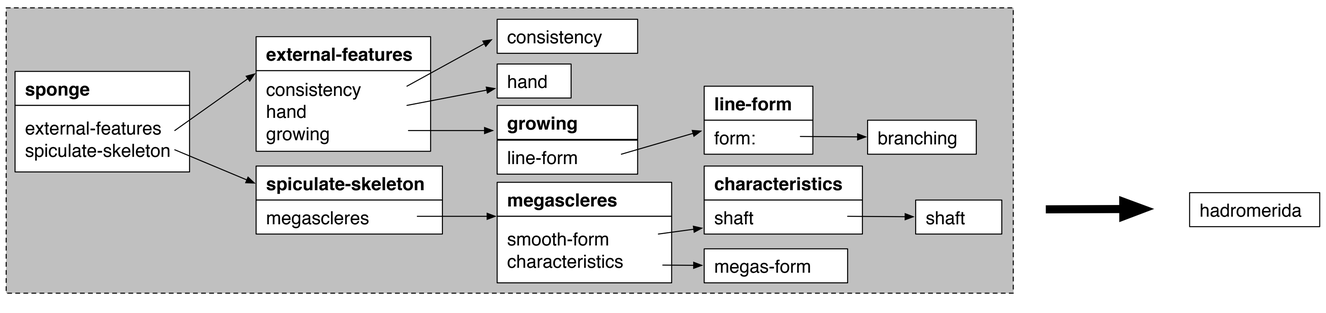
Concept Convergence

$$R_1^{10} = \{\alpha_1, \alpha_2, \alpha_3, \alpha_5, \alpha_6, \alpha_9\}$$
$$R_2^{10} = \{\beta_1, \beta_2, \beta_3, \beta_5\}$$

$t = 0$	$A_1 : \alpha_1, ..., \alpha_6$
	$A_2 : \beta_1, ..., \beta_5$
$t = 1$	$A_2 : \beta_6 \twoheadrightarrow \alpha_4$
$t = 2$	$A_1 : \alpha_7 \twoheadrightarrow \beta_6$
$t = 3$	$A_2 : \beta_7 \twoheadrightarrow \alpha_7$
$t = 4$	$A_1 : retract(\alpha_4)$
	$A_1 : \alpha_8$
$t = 5$	$A_2 : \beta_8 \twoheadrightarrow \alpha_8$
$t = 6$	$A_1 : e_{251} \twoheadrightarrow \beta_8$
$t = 7$	$A_2 : e_{26} \twoheadrightarrow \alpha_8$
$t = 8$	$A_1 : retract(\alpha_8)$
	$A_1 : \alpha_9$



Example Argument



Belief Revision

When agents use A-MAIL:
- They generate initial hypotheses and share them
- Agents attack each others hypotheses sending arguments
- When an agent receives an argument, it might change its beliefs

When an agent receives arguments:

1) If it is an example: add to local set of examples
2) Re-assess which arguments are defeated
3) If any positive examples become uncovered, use ABUI to generate new rules that cover them