Learning Social Integrity Constraints

Fabrizio Riguzzi
University of Ferrara

Introduction

- A theory composed of social integrity constraint can be seen as a classifier
- It classifies an history as compliant or non-compliant
- Similar to the learning from interpretation setting of Inductive Logic Programming

Learning from Interpretation

- Aim: learning a classifier for logical interpretations
- Classifier: a set of disjunctive clauses
- Disjunctive clause $C = h_1 \vee h_2 \vee \ldots \vee h_n \leftarrow b_1, b_2, \ldots, b_m$
- $head(C) = \{h_1, h_2, \ldots, h_n\}$
- $body(C) = \{b_1, b_2, \ldots, b_m\}$
- Interpretation = set of ground atoms.

Learning from Interpretation

- Set of clauses as a classifier
  - an interpretation is positive if all the clauses are true in the interpretation
  - an interpretation is negative if there exists at least one clause that is false in it
- A clause C is true in an interpretation I if for all grounding substitutions θ of C:
  $I \models body(C)\theta \rightarrow head(C)\theta \cap I \neq \emptyset$.

Test of Truth of a Clause

- Range restricted clause C, finite interpretation I: run the query $?-body(C), not\ head(C)$ against a logic program containing I
- If $C = h_1 \vee h_2 \vee \ldots \vee h_n \leftarrow b_1, b_2, \ldots, b_m$ then the query is $?-b_1, b_2, \ldots, b_m, not\ h_1, not\ h_2, \ldots, not\ h_n$
- If the query succeeds, C is false in I. If the query fails, C is true in I

Example

- $I = \{invitation(a), accept(a), blood\_test(a), invitation(b), refusal(b)\}$
- $C = accept(X) \vee refusal(X) \leftarrow invitation(X)$: the clause is true in I because the query $? - invitation(X), not\ accept(X), not\ refusal(X)$ fails
- $C = blood\_test(X) \leftarrow invitation(X)$: the clause is false in I because the query $? - invitation(X), not\ blood\_test(X)$ succeeds with $\theta = \{X/b\}$.

Learning from Interpretations

Given

- a space of possible clausal theories $\mathcal{H}$
- a set P of interpretations
- a set N of interpretations

Find: a clausal theory $H \in \mathcal{H}$ such that

- for all $p \in P$, H is true in p
- for all $n \in N$, H is false in n

Application to ICs Learning

- Interpretations=histories
- Clause theory=ICs

Example from guidelines discovery

- Each history (trace, guideline execution) is an interpretation
- Cervical cancer screening process
  $\{\mathbf{H}(invitation, 1), \mathbf{H}(papTest, 2), \mathbf{H}(sendPapTestSample, 3),$
  $\mathbf{H}(sendPapTestResult(neg), 4),$
  $\mathbf{H}(sendNegativeLetter, 5)\ \}$
- clauses $\rightarrow$ ICs

$$\mathbf{H}(invitation, T) \rightarrow$$

$$\mathbf{E}(papTest, T1) \wedge T1 > T \vee \mathbf{E}(refusal, T2) \wedge T2 > T$$

$$\mathbf{H}(sendPapTestResult(neg), T) \rightarrow \mathbf{EN}(papTest, T1) \wedge T1 > T$$

Learning Algorithm

- ICL [De Raedt, Van Laer, 95]

Learn(P, N, B)
$H := \emptyset$
repeat until best clause C not found or N is empty
    find best clause C
    if best clause C found then
        add C to H
        remove from N all interpretations that are false for C
return H

Find best clause

- Find best clause performs a beam search in the space of clauses ordered according to the θ-subsumption generality order
- Clause C θ-subsumes clause D is $\exists\theta$ such that $C\theta \subseteq D$
- if D θ-subsumes C than clause C is more general than clause D (is true in more interpretations)
- In particular if $D \subseteq C$ than clause C is more general than clause D (is true in more interpretations)
- Example
  $C = accept(X) \vee refusal(X) \leftarrow invitation(X)$
  $D = accept(X) \vee refusal(X) \leftarrow true$
  $D = accept(X) \leftarrow invitation(X)$

Find best clause

- Find best clause starts from the most specific clause $\{\} = false \leftarrow true$ and gradually generalizes it until all the positive interpretations are covered and some negative interpretations are ruled out
- Generalization is performed by adding literals to the clause (to the head or to the body)
- The literals that can be added are specified using the language bias
- Example:
  $false \leftarrow true$
  $accept(X) \leftarrow true$
  $accept(X) \leftarrow invitation(X)$
  $accept(X) \vee refusal(X) \leftarrow invitation(X)$

ICs learning

- In order to learn ICs we must find a way to generalize them
- Adding a literal to the body makes the IC true in more histories

$$\mathbf{H}(invitation, T) \rightarrow \mathbf{E}(papTest, T1) \wedge T1 > T$$

$$\mathbf{H}(invitation, T) \wedge \mathbf{H}(accept, T3) \rightarrow$$
$$\mathbf{E}(papTes, T1) \wedge T1 > T$$

- Adding a disjunct to the head makes the IC true in more histories

$$\mathbf{H}(invitation, T) \wedge \mathbf{H}(accept, T3) \rightarrow$$
$$\mathbf{E}(papTes, T1) \wedge T1 > T \vee \mathbf{E}(refusal, T2) \wedge T2 > T$$

Generalizing ICs

- Adding a conjunct to a disjunct with universal quantification (disjunct with $\mathbf{EN}$) in the head generalizes the IC
  $\mathbf{H}(sendPapTestResult(neg), T) \rightarrow \mathbf{EN}(papTest, T1)$
  $\mathbf{H}(sendPapTestResult(neg), T) \rightarrow$
  $\mathbf{EN}(papTest, T1) \wedge T1 > T$
- Adding a conjunct to a disjunct with existential quantified variables (disjunct with $\mathbf{E}$) in the head specializes the IC
  $\mathbf{H}(invitation, T) \rightarrow \mathbf{E}(papTest, T1) \vee \mathbf{E}(refusal, T2)$
  $\mathbf{H}(invitation, T) \rightarrow$
  $\mathbf{E}(papTest, T1) \vee \mathbf{E}(refusal, T2) \wedge T2 > T$

Generalization Operator

- add a literal to the body
- add a disjunct to the head composed of:
  - a single $\mathbf{EN}$ atom
  - an $\mathbf{E}$ atom followed by all the constraints allowed by the language bias for the disjunct
- add a constraint to a $\mathbf{EN}$ disjunct
- remove a constraint from a $\mathbf{E}$ disjunct

Experiments

- Cervical cancer screening process
- 344 positive histories, 525 negative histories
- 15 possible events:
  $invitation, papTest, sendPapTestSample,$
  $sendPapTestResult(neg), sendPapTestResult(pos),$
  $sendNegativeLetter, colposcopy, biopsy,$
  $sendBiopsySample, sendBiopsyResult(neg),$
  $sendBiopsyResult(pos), positivePhoneCall, refusal,$
  $sendColposcopyResult(neg),$
  $sendColposcopyResult(dubious)$

Cervical Cancer Screening

- Language bias:
  - atoms $\mathbf{H}(e, T)$ and $\mathbf{H}(e, T1)$ in the body for every possible event e where T and $T1$ are two variables
  - a positive disjunct in the head for every possible event e containing the expectation $\mathbf{E}(e, T2)$ and the constraints
    $T < T2 - 1, T1 < T2 - 1, T = T2 - 1, T1 = T2 - 1$
  - a negative disjunct in the head for every possible event e containing the expectation $\mathbf{EN}(e, T2)$ and the constraints
    $T < T2 - 1, T1 < T2 - 1, T = T2 - 1, T1 = T2 - 1$
- Results: a complete and consistent theory, learned in 40 minutes

Theory Learned

7 ICs, among which:

$\mathbf{H}(biopsy, T)) \land \mathbf{H}(sendPapTestResult(neg, T1) \rightarrow false.$

Excludes 8 negatives

$\mathbf{H}(biopsy, T) \land \mathbf{H}(invitation, T1) \rightarrow$
$\mathbf{EN}(biopsy, T2) \land T1 = T2 - 1.$

Excludes 25 negatives

$\mathbf{H}(sendPapTestResult(pos, T) \land \mathbf{H}(invitation, T1)) \rightarrow$
$\mathbf{E}(colposcopy, T2) \land T < T2 - 1 \land T1 < T2 - 1.$

Excludes 9 negatives

$\mathbf{H}(invitation, T) \rightarrow$
$\mathbf{E}(papTest, T1) \land T = T1 - 1 \lor \quad \mathbf{E}(refusal, T1) \land T = T1 - 1.$

Excludes 456 negatives

$\mathbf{H}(sendPapTestResult(neg), T) \land \mathbf{H}(invito, T1) \rightarrow$
$\mathbf{EN}(papTest, T2) \land T < T2 - 1.$

Excludes 25 negatives

Experiments

- Combinatorial auction
- 324 positive histories, 3676 negative histories randomly generated
- Example IC:

$$\begin{aligned}
&\mathbf{H}(tell(a,b,op_au(item,TEnd,TDead),auction1),TOpen)\wedge\\
&\mathbf{H}(tell(b,a,bid(item,Quote),auction1),TBid)\rightarrow\\
&\mathbf{E}(tell(a,b,answer(lose,item,Quote),auction1),TLose)\wedge\\
&\quad TLose \leq TDead \wedge TEnd < TLose \vee\\
&\mathbf{E}(tell(a,b,answer(win,item,Quote),auction1),TWin)\wedge\\
&\quad TWin \leq TDead \wedge TEnd < TWin
\end{aligned}$$

Future Works

- More experiments
- Application of other learning from interpretation algorithms: Claudien, WARMR
- Improvement of the bias: at the moment just sets, in the future more elaborate