Relational Learning from Drug Adverse Events Reports

Ergun M. Biçici

North Carolina State University

October 14, 2004

*This work was supported in part by GlaxoSmithKline. The author would like to thank Gary H. Merrill and Alan J. Menius of GlaxoSmithKline Data Exploration Sciences for helpful discussions and for their guidance and support during the term of this research.

Outline

- Problem: Learning from adverse event reports
- Approach: Discover rules using relational learning with FOIL
- Experiments: Adverse events data, setup
- Discussion of results

Mining Adverse Event Reports

- *A drug adverse event is any unintended response in a patient's body during or after the use of a drug.*
- Reports of adverse events contain valuable information for both the FDA [1] and the pharmaceutical industry.
- Machine learning challenge: discovering rules from a data set that spans over multiple domains.

Drug Adverse Event Reports

- An adverse event report may contain information about:
  - the dose and the frequency of intake (e.g. 10mg, daily)
  - the demographics of the patient (e.g. age, weight, gender)
  - the set of adverse reactions observed (e.g. eye infection, difficulty in breathing), and
  - the concomitant medication (e.g. aspirin, antibiotics)

- We used reports pertaining to the drugs withdrawn from the market:
  - more likely to contain information that is of value in understanding the causal links
  - the causal links already established by the FDA for withdrawal provides us a way to evaluate the effectiveness of our approach.

Sample Adverse Event Report

Detailed Results for BAYCOL – Report Id 141580

Drug Involvement	Suspected
Date Report Received at MHPD	2001-08-16
Outcome	Recovered without sequelae
Serious	Y
Age at Reaction	73
Gender	F
Weight	68
Feature of Report	Adverse reaction
Type of Report	Spontaneous
Route of Administration	Per oral
Amount	.8 Milligrams
Frequency	1 Daily
Dose Duration	29 Day(s)
Adverse Reaction Term	• CPK INCREASED • MUSCLE PAIN • RHABDOMYOLYSIS • WALKING DIFFICULTY • WEAKNESS GENERALIZED
Other Drugs Involved	• ALTACE(Concomitant) • ALTACE(Concomitant) • PREMARIN(Concomitant) • PREMARIN(Concomitant) • PLAVIX(Concomitant) • NADOLOL(Concomitant) • DIGOXIN(Concomitant) • LASIX(Concomitant) • NADOLOL(Concomitant)
Notifiers	• Physician/Manufacturer

Drug Adverse Events Data

- From Health Canada's Canadian Adverse Drug Reaction Information System (CADRIS) [2].

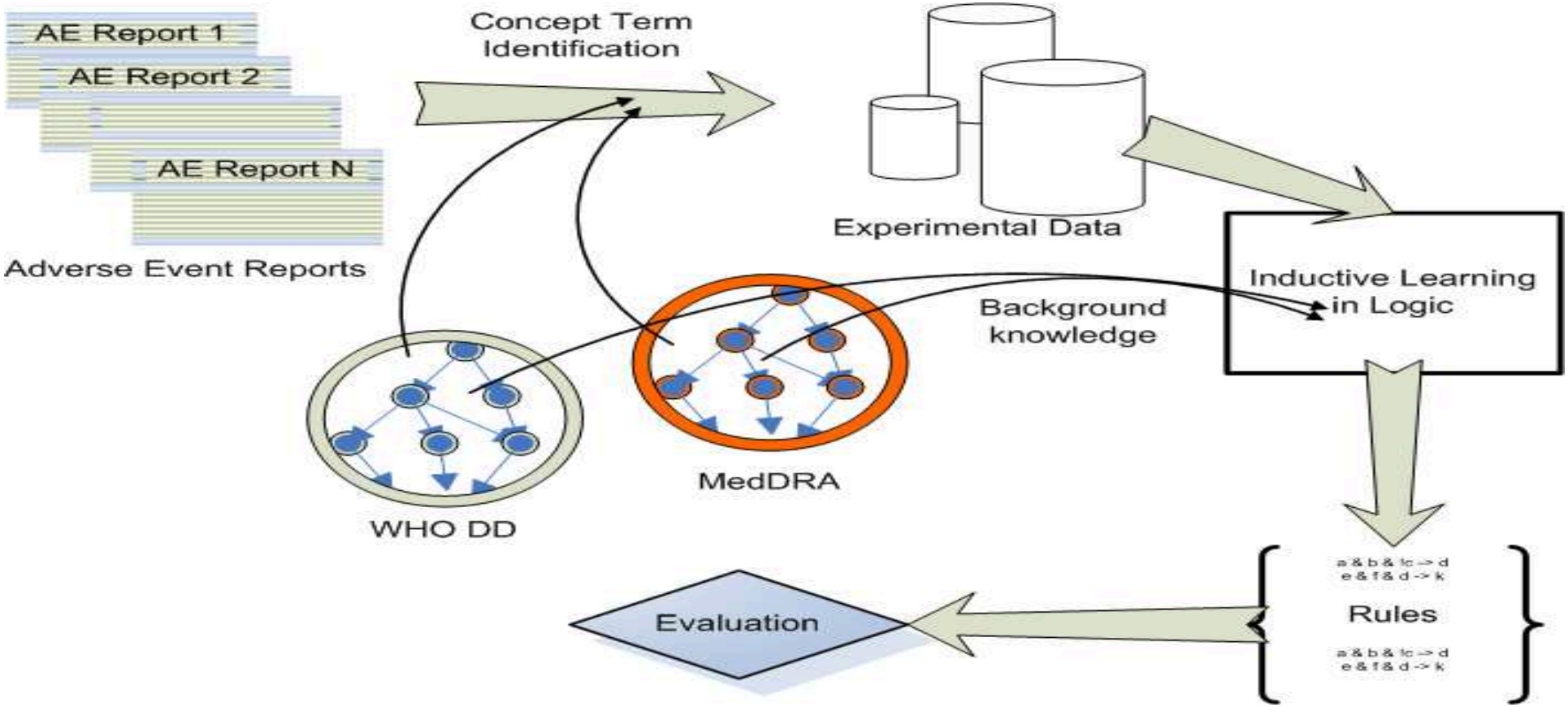
- Produced a data set for each drug that contains tuples of:
  - $Drug\text{-}AdverseEvent\text{-}ConcomitantMedication$
  - $Drug\text{-}AdverseEvent\text{-}ConcomitantMedication\text{-}SeriousEffect\text{-}Gender$
  - $Drug\text{-}AdverseEvent\text{-}ConcomitantMedication\text{-}SeriousEffect\text{-}Age\text{-}Gender\text{-}Weight$

  which enter the learning as arguments to the relation "$associatedWith$" depending on the experiment.

- Used domain ontologies for background knowledge:
  - WHO Drug Dictionary [4] for information about drugs and
  - MedDRA [3] for information about adverse events

- Used Babylon Knowledge Explorer [5] to access ontologies.

System Overview



Overview of the learning system.

Relational Learning

- Multiple sources of information are related to each other within the context of each report.

- Our learning task was:

  To find a classification of the adverse events each drug is associated with, with or without the presence of other medication and with or without the presence of other information including the gender, age, weight, and the seriousness of the adverse event.

- From our experiments, we expect to find rules of the form:

  $$associatedWith(Pondimin, A, C) \leftarrow$$
  $$isaA(A, HeartDamage), isaO(C, Diuretics),$$ [1]

[1] The $\leftarrow$ sign represents the implication operator in which the head appears on the left hand side and the body on the right hand side. The "," sign "and"s the relations specified in the body.

First Order Inductive Learning (FOIL)

- An efficient and widely used top-down inductive learning system that can learn function-free Horn clauses.
- Uses information-gain metric to generate rules
- Background relations used to represent reports:
  - $associatedWith(Drug, AdverseEvent, ConcomitantMedication)$: $Drug$ is associated with $AdverseEvent$ when used with $ConcomitantMedication$.
  - $isaD(Drug1, Drug2)$: $Drug1$ "isa" type of $Drug2$.
  - $isaA(AdverseEvent1, AdverseEvent2)$: $AdverseEvent1$ "isa" type of $AdverseEvent2$.
  - $isaO(ConcomitantMedication1, ConcomitantMedication2)$: $ConcomitantMe$ "isa" type of $ConcomitantMedication2$.

Experiments

- The reports were taken from the following set of drugs: CISAPRIDE, HISMANAL, PONDIMIN, RAXAR, REDUX, REZULIN, and SELDANE.
- We used FOIL to learn:
  - rules per withdrawn drug
  - rules for all reports of withdrawn drugs
- Explored the change in the accuracy and coverage when different sources of information are incorporated.

Some Rules Generated

- $associatedWith(Pondimin, B) \leftarrow$
  $isaA(B, C),\ isaA(C, D),\ isaA(D, Renal\ and\ Urinary\ Disorders).$

  $Pondimin$ is associated with an adverse event which is a type of renal and urinary disorder.

- $associatedWith(Hismanal, Hypertension, C, D) \leftarrow C <> D.$

  $Hismanal$ is associated with hypertension when the attributes for seriousness of the report and the gender are not equivalent.

- $associatedWith(Rezulin, B, Serious, Female) \leftarrow$
  $associatedWith(Rezulin, B, NotSerious, Male),$
  $B <> HAEMOGLOBIN\ DECREASED.$

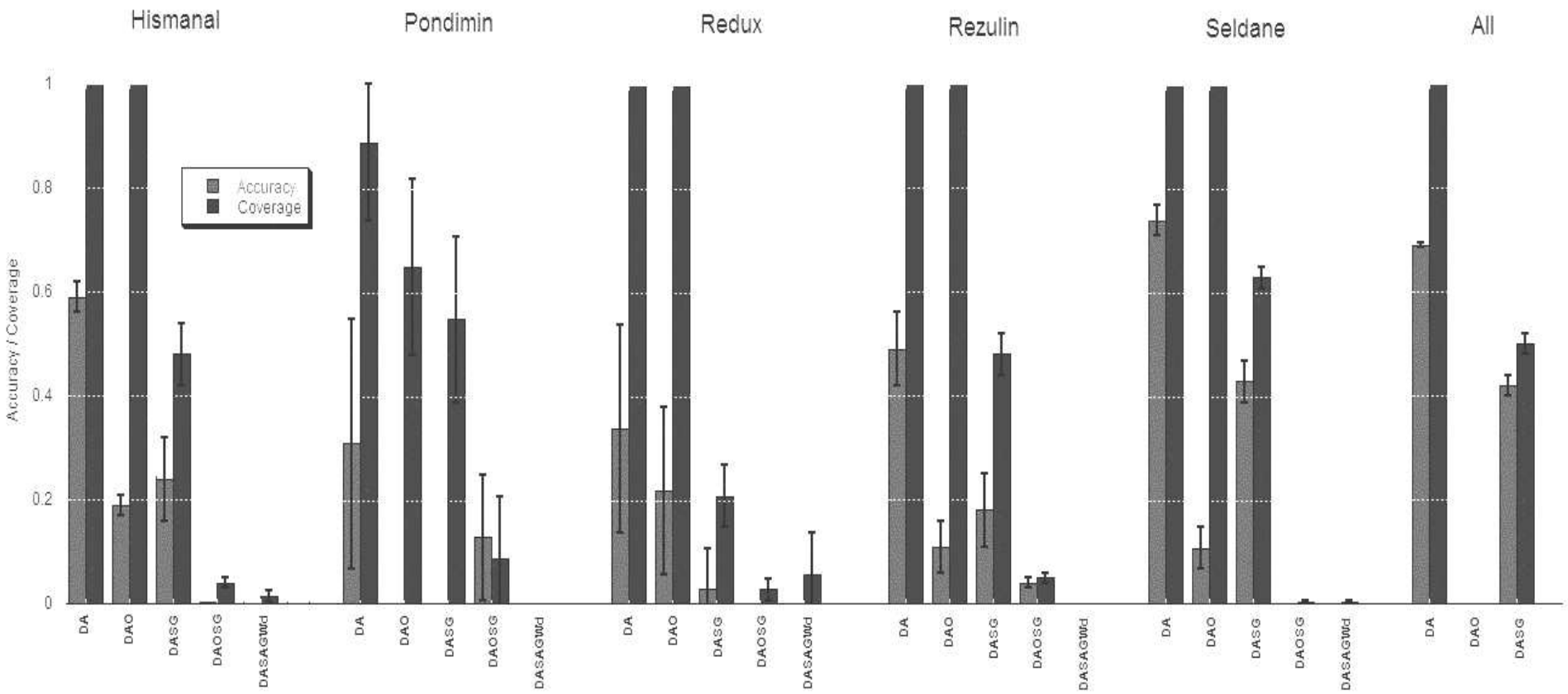
Results (2)

$associatedWith(Hismanal, ECG\ abnormal, Lidocaine, Y, N)$
$associatedWith(Hismanal, Dyspnoea, Oxygen, Y, N)$
$associatedWith(Hismanal, QT\ prolonged, Licodaine, Y, N)$
$associatedWith(Hismanal, Cardiac\ arrest, Licodaine, Y, N)$
$associatedWith(Hismanal, Arrhythmia\ ventricular, Licodaine, Y, N)$

Table 1: Some rules induced by FOIL when learning from $Hismanal$ data with DAOSG attribute set.

$Hismanal$ was withdrawn due to cardiac arrest and arrhythmias.

Results (3)



Accuracy-coverage of the rules learned by FOIL corresponding to the withdrawn drugs with increasing sources of information included from the adverse events reports.

Discussion

- We show the feasibility of the rules generated by relational learning in explaining the adverse event reports.
- Explanations can match the reasons for drug withdrawals.
- We can discover specific but highly accurate rules.

References

[1] Food and Drug Administration. URL: http://www.fda.gov.

[2] Health Canada's Canadian Adverse Drug Reaction Information System (CADRIS). URL: http://www.hc-sc.gc.ca/hpfb-dgpsa/tpd-dpt/.

[3] MedDRA: Medical Dictionary for Regulatory Activities. URL: http://www.meddramsso.com.

[4] The World Health Organization Drug Dictionary (WHO DD). http://www.who-umc.org/faqs/faqdd.html.

[5] Gary H. Merrill. The Babylon Project: Toward an extensible text-mining platform. *IEEE IT Professional*, 5(2):23–30, 2003.

[6] J. R. Quinlan and R. M. Cameron-Jones. FOIL: A midterm report. In P. B Brazdil, editor, *Machine Learning (ECML-93) European Conference on Machine Learning Proceedings; Vienna, Austria*, pages 3–20, Berlin, Germany, 1993. Springer-Verlag.

Thank you!