Imperial College London

Ultra-Strong Machine Learning Comprehensibility of Programs Learned with Inductive Logic Programming

Stephen Muggleton, Ute Schmid, Christina Zeller, Alireza Tamaddoni-Nezhad, Tarek Besold

Department of Computing
Imperial College, London, UK

Motivation

- Michie (1988) - definition of Ultra-Strong Machine Learning requires a) predictive accuracy increase, b) hypotheses in symbolic form and c) human performance increase after study of machine-generated hypotheses
- Mitchell (1997) - definition of Machine Learning in terms of Predictive Accuracy alone
- ILP and symbolic AI generally need operational definition of comprehensibility to distinguish communicable and non-communicable knowledge
- Testability in age of Mechanical Turk

Text comprehension tests

"For many years people believed the cleverest animals after man were chimpanzees. Now, however, there is proof that dolphins may be even cleverer than these big apes."

Question: Which animals do people think may be the cleverest?

[http://englishteststore.net]

Program comprehension tests

```
p(X,Y) :- p1(X,Z), p1(Z,Y).
p1(X,Y) :- father(X,Y).
p1(X,Y) :- mother(X,Y).
father(john,mary). mother(mary,harry).
```

Question: p(john,harry)?

Initial experiments - recognisable predicates

```
p(X,Y)  :- p1(X,U), p1(U,Z), p1(Z,Y).
p1(X,Y) :- father(X,Y).
p1(X,Y) :- mother(X,Y).
```

parent

grandgrand parent

Tentative finding: Annotation strategy appears to beat tabulation and manual inference.

More recent experiment - chemistry domain

Background	Example	Target
q1(ab,ac)	exo(ac,an)	exo(X,Y) :- q1(X,Z), q1(Z,Y)
q2(aa,ac)	not exo(aa,ab)	exo(X,Y) :- q1(X,Z), q2(Z,Y)
q1(ad,ag)	exo(ab,ag)	exo(X,Y) :- q2(X,Z), q2(Z,Y)
q2(ad,ae)	not exo(ad,ai)	exo(X,Y) :- q2(X,Z), q1(Z,Y)
. . .	. . .	

Definitions

- Comprehensibility - proportion of correct answers after inspection of program [C]
- Inspection time [T] - time taken to read program
- Predicate recognition [R] - mean proportion predicates correctly recognised
- Naming time [N] - time to name predicate
- Textual complexity [Sz] - program size
- Unaided Human Comprehension of Examples C(S,E)
- Machine-aided Human Comprehension of Examples C(S,M(E))

Experimental Hypotheses

H1 $C \propto \frac{1}{T}$ - long inspection time related to incomprehension

H2 $C \propto R$ - comprehension related to recognition of predicate

H3 $C \propto \frac{1}{Sz}$ - long programs hard to understand

H4 $R \propto \frac{1}{N}$ - long naming time related to lack of recognition

H5 $C(S, E) < C(S, M(E))$ - improved human performance after studying machine-learned rules

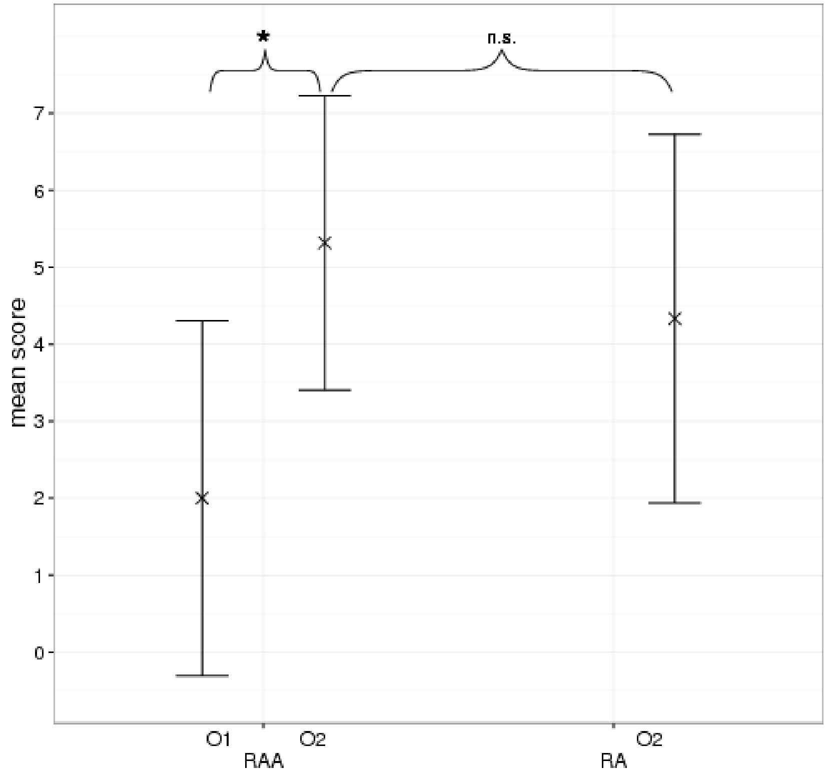
Experiment participants

Participants were undergraduate students of cognitive science (20 female, 23 male, mean age = 22.12 years, sd = 2.51) with a good background in Prolog.

Experimental Results - Family Relations

H1	Statistically confirmed
H2	Statistically confirmed
H3	Partially confirmed
H4	Partially confirmed - recursive ancestor exception
H5	Statistically confirmed

H5 result



Mean comprehensibility scores for rule acquisition and application (RAA) vs. rule application (RA)

Conclusions and further work

- First operational definition of comprehensibility
- First demonstration of Michie's Ultra-Strong Machine Learning
- Confirmation of hypotheses
- Difficulties in understanding recursion- eg ancestor/2
- Value of operational definition of comprehension to AI systems development
- A theory of the Explainable

Bibliography

- D. Michie. Machine learning in the next five years. In Proceedings of the Third European Working Session on Learning, pages 107122. Pitman, 1988.
- U. Schmid, C. Zeller, T. Besold, A. Tamaddoni-Nezhad, and S.H. Muggleton. How does predicate invention affect human comprehensibility?. In Alessandra Russo and James Cussens, editors, Proceedings of the 26th International Conference on Inductive Logic Programming, pages 52-67, Berlin, 2017. Springer-Verlag.
- S.H. Muggleton, U. Schmid, C. Zeller, A. Tamaddoni-Nezhad, and T. Besold. Ultra-strong machine learning - comprehensibility of programs learned with ILP. Machine Learning, 107:1097-1118, 2018.