Developing expert *in silico* systems

Lhasa Limited

Director of Science
chris.barber@lhasalimited.org

Chris Barber

Developing expert *in silico* systems

- What is a knowledge-base expert system?

- How should they be used?

A knowledge-based expert system uses human knowledge to solve problems that normally would require human intelligence

- “An expert system is one that gives the answers an expert would give…

  …including the wrong ones”. P. Judson

- Emulates the decision-making ability of a human expert
  - Encodes (expert) human knowledge and reasoning
    - This was the origin of Derek (Deductive Estimation of Risk from Existing Knowledge)
      - …to capture the (toxicity) knowledge of experts
  - Rule-based expert systems process the ‘rules’ as encoded
  - Knowledge-based expert systems apply reasoning to a knowledge base of facts
    - An inference engine reasons with facts from the knowledge-base.

A Review on Knowledge-based Expert System: Concept and Architecture, Tripathi.. *Artificial Intelligence Techniques, 2011, 19*
Knowledge-Based Expert Systems in Chemistry: Not Counting on Computers, Judson. book - RSC Theoretical and Computational Chemistry Series, Sep 2009

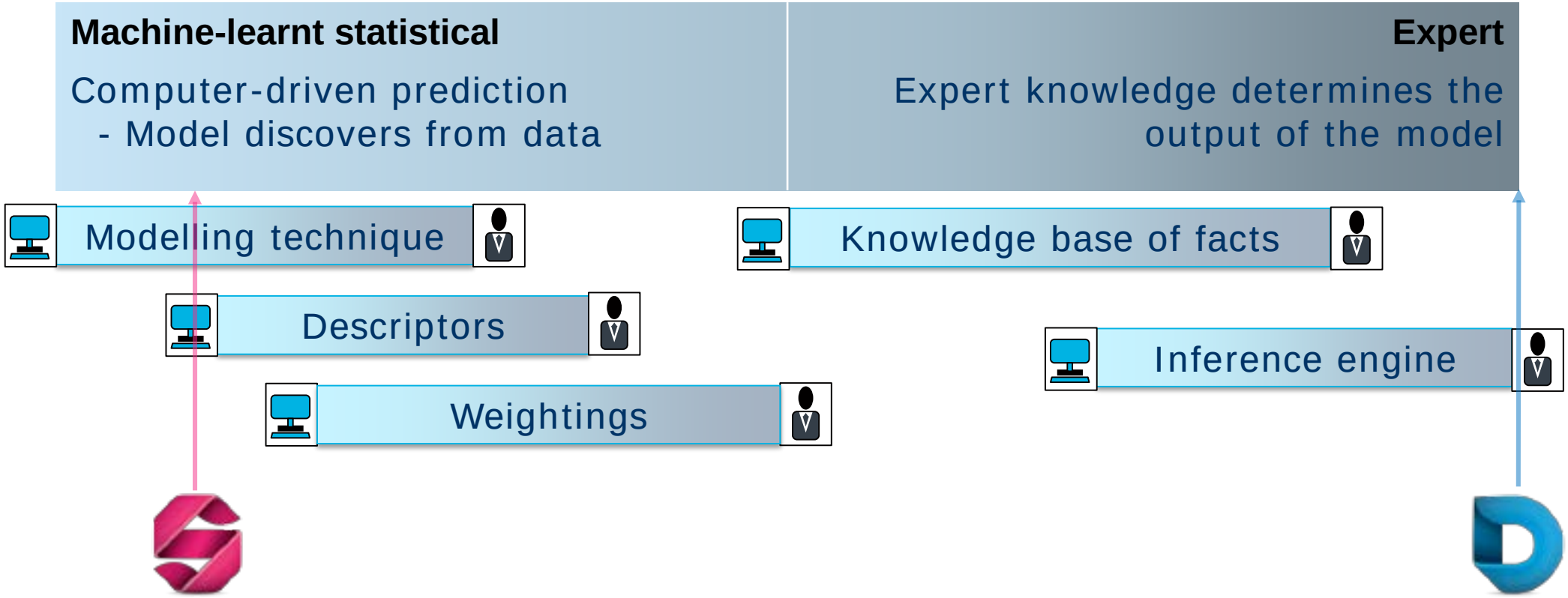
A knowledge-based expert system uses human knowledge to solve problems that normally would require human intelligence

- Simple rule-based systems do not deal with uncertainty
- Knowledge-based systems can deal with uncertainty

Terms of uncertainty (Derek Nexus)	Definition
Certain	There is proof that the proposition is true
Probable	There is at least one strong argument that the proposition is true and there are no arguments against it
Plausible	On balance the weight of evidence supports the proposition
Equivocal	There is an equal weight of evidence for and against the proposition
Contradicted	There is proof both that the proposition is true and that it is false

A Review on Knowledge-based Expert System: Concept and Architecture, Tripathi.. *Artificial Intelligence Techniques, 2011, 19*
Using argumentation for absolute reasoning about the potential toxicity of chemicals, Judson.., *J. Chem. Inf. Comp. Sci.*, 2003, 1364

Differences between (knowledge-base) expert and statistical systems…



Machine learnt

Expert defined

- There is no clear boundary between expert and statistical models
  - Injection of human knowledge *vs.* how much the model ‘discovers’?
- The ‘right’ balance depends upon the use-case

Developing expert *in silico* systems

- What is a knowledge-base expert system?

- How should they be used?

Applicability domains - superficially a good idea but it is not enough

- “Is my compound in the applicability domain?” requires the modeller to know…
  - The use-case and required level of accuracy
  - The consequences of a mis-prediction
  - Knowledge of everything a user may put into the model!
- Many attempts to define applicability domains are fundamentally flawed…
  - Comparison of different approaches to define the applicability domain of QSAR models. Sahigara, Molecules, 2012, *17*(5), 4791
- Applicability domain is a ‘catch-all’ term that confuses
  - Appropriateness of the model *for the prediction*
  - Relevance, quantity and quality of supporting data used *for the prediction*
  - Uncertainty in the model *for the prediction*
- ‘Applicability domain’ attempts to hide the need for a user to understand, review and make an expert assessment on the prediction…

The user should be asking “how much can I trust this prediction?”

- ..and should expect to receive
  - An explanation for the prediction
    - What knowledge is behind this prediction?; What is the reasoning?; The mechanism?
  - Supporting data
    - What facts are known?; How relevant are they to my compound?; How certain are they?
  - A measure of reliability
    - What is the uncertainty in the prediction?
    - How accurate is the model for compounds like mine?
    - What aspects give the greatest risk to the prediction?

- You wouldn’t blindly accept the judgement of a self-declared human expert or an in vivo study without asking questions about the expertise / protocol…
  - Why would you unquestioningly trust an *in silico* model?

See also Thierry’s talk earlier this morning Applicability domains: towards a more formal definition

Our approach to answer the question "how much can I trust this prediction?" …as implemented in Derek for mutagenicity

- Emulates a conversation with a human expert..
- "I'm predicting negative as I see no reasons for activity (alerts)..
  - … and I'm very experienced; all the features of this compound have been seen in negative compounds. Nothing worries me!"
  - … but I have seen this feature in a false-negative prediction – you may want to review that part.."

misclassified feature

  - … but that bit is unusual for me – you may want to review it.."

unclassified feature

It's difficult, but important, to make negative predictions Williams… *Reg. Tox. and Pharmacol.* 2016, *76*, 79
Establishing best practise in the application of expert review of mutagenicity under ICH M7 Barber... *Reg. Tox. and Pharmacol.* 2015, *73*, 367

Summary

- What is an knowledge-base expert system?
  - What is special about them?
  - What distinguishes these from statistical models?

- How should they be used?
  - The need for expert assessment of the output
  - Moving from applicability domains to presenting information that supports informed decision-making

THE QUEEN'S AWARDS FOR ENTERPRISE: INNOVATION 2016

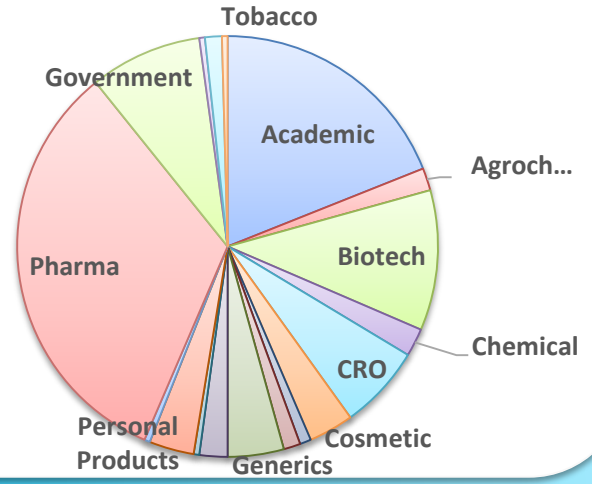
Lhasa has just received the prestigious **Queen's Award for Enterprise 2016, Innovation** in recognition of the development of the expert knowledge-based software system, Derek Nexus.

Not-for-profit Educational charity

- Teaching lectures
- Interns
- Sponsor PhDs
- Undergrad projects

Located in Leeds, UK

A membership organisation



Lhasa Limited

Predictive software
(expert & machine learnt)

Purge

Degradation

Toxicity

Metabolism

Regulators
(members & collaborators)

FDA

NIHS

OECD
BETTER POLICIES FOR BETTER LIVES

Data & knowledge sharing Honest broker

Vitic intermediates

Vitic excipients

eTOX

MIP-DILI

iPiE

Proprietary data mining

Lhasa Limited

shared **knowledge** • shared **progress**

Lhasa Limited Registered Office
Granary Wharf House, 2 Canal Wharf, Leeds LS11 5PS
Registered Charity (290866)

+44 (0)113 394 6020
info@lhasalimited.org
www.lhasalimited.org

Company Registration Number 01765239. Registered in England and Wales. VAT Registration Number GB 396 8737 77.

ISO 9001: CERTIFIED

Abstract

- **Developing expert *in silico* systems.**

- The application of QSAR for the prediction of chemical properties has matured significantly during the past 10 years driven by a number of factors including size and availability of data, improvements in machine-learning techniques and the pressure to reduce *in vitro* and *in vivo* testing for reasons of time, cost and the use of animals in testing.
- This talk will address key questions including ‘what distinguishes an expert system from a statistical one?’, ‘what makes a good expert system?’, ‘how should applicability domains be defined for expert systems?’, ‘how can machine learning enhance an expert system?’ and will be illustrated with recent advances in the expert systems Derek Nexus and Meteor Nexus.