CITY UNIVERSITY LONDON

Neural-Symbolic Learning and Reasoning

DAGSTUHL SEMINAR 14381, SEPTEMBER 2014

ARTUR D'AVILA GARCEZ

Neural-Symbolic Learning and Reasoning: Contributions and Challenges

Artur d'Avila Garcez[1], Tarek R. Besold[2], Luc de Raedt[3], Peter Földiak[4], Pascal Hitzler[5], Thomas Icard[6], Kai-Uwe Kühnberger[2], Luis C. Lamb[7], Risto Miikkulainen[8], Daniel L. Silver[9]

Knowledge representation: computer science logic

Consolidation: knowledge extraction and transfer learning

Killer app: big data + descriptions
(e.g. reasoning about video and audio)

Neural-Symbolic Learning and Reasoning Association: www.neural-symbolic.org

Representation precedes Learning

We need a language for describing the alternative algorithms that a network of neurons may be implementing…

Computer Science Logic + Neural Computation

GOAL of NSI: Learning from experience and reasoning about what has been learned from an uncertain environment in a computationally efficient way

Les Valiant

History of neural-symbolic integration (1)

1988: P Smolensky, On the proper treatment of connectionism, BBS:11(1); J McCarthy (commentary), Epistemological challenges for connectionism

1990: G Hinton, Preface to the special issue on connectionist symbol processing, Artificial Intelligence 46,1-4

1993: L Shastri and V Ajjanagadde, From simple associations to systematic reasoning: A connectionist representation of rules, variables and dynamic bindings using temporal synchrony, BBS:16 (SHRUTI)

1994: G Towell, J Shavlik, Knowledge-Based Artificial Neural Networks. Artif. Intell. 70(1-2): 119-165 (KBANN)

1997: M Craven, J Shavlik, Understanding Time-Series Networks: A Case Study in Rule Extraction. Int. J. Neural Syst. 8(4): 373-384 (TREPAN)

2001: A Browne, R Sun, Connectionist inference models. Neural Networks 14(10): 1331-1355

2002: A d'Avila Garcez, K. Broda and D Gabbay, Neural-Symbolic Learning Systems: Foundations and Applications, Perspectives in Neural Computing, Springer-Verlag. (CILP)

History of neural-symbolic integration (2)

2006: A d'Avila Garcez, L Lamb, A Connectionist Computational Model for Epistemic and Temporal Reasoning. Neural Computation 18(7): 1711-1738

2007: S Bader, P Hitzler, S Hölldobler, A. Witzel, A Fully Connectionist Model Generator for Covered First-Order Logic Programs. IJCAI 2007: 666-671

2007: Y Bengio, Y LeCun, Scaling learning algorithms towards AI. Large-scale kernel machines 34(5) (Representation Learning)

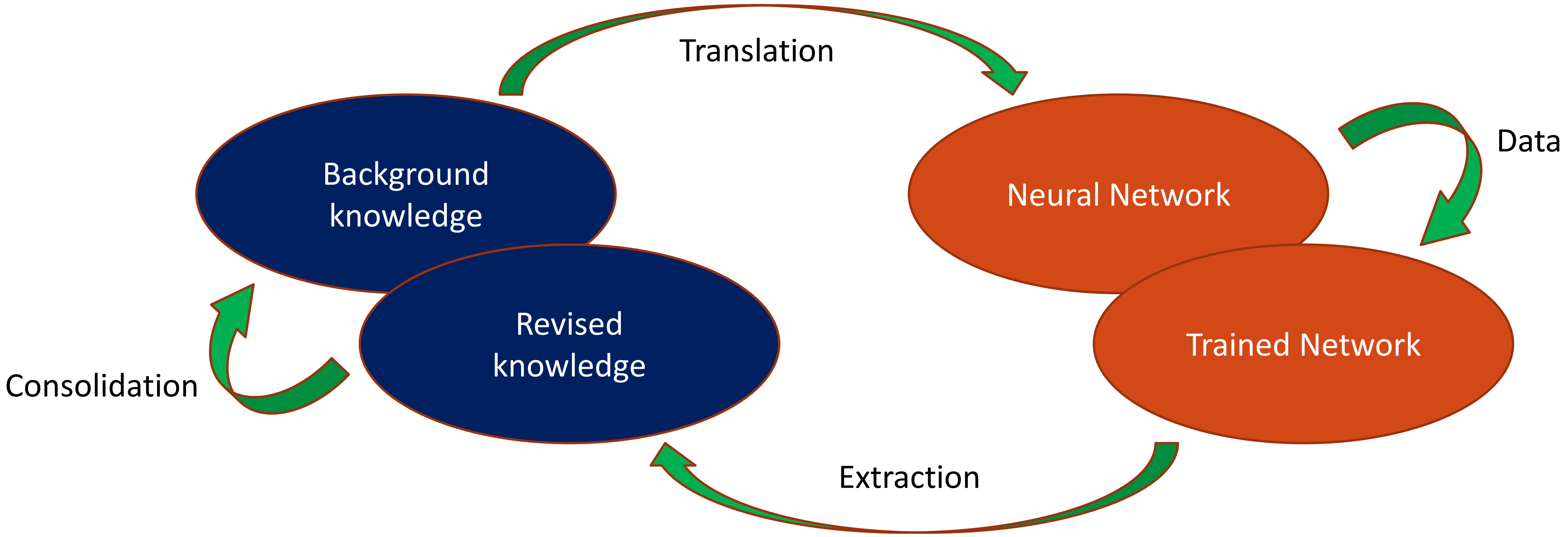
2009: A d'Avila Garcez, L Lamb, D. Gabbay, Neural-Symbolic Cognitive Reasoning. Cognitive Technologies, Springer

2010: D. Endres, P. Foldiak, U. Priss, An application of formal concept analysis to semantic neural decoding. Annals of Mathematics and Artificial Intelligence 57(3-4): 233-248

2012: G Pinkas, P Lima, S Cohen, A Dynamic Binding Mechanism for Retrieving and Unifying Complex Predicate-Logic Knowledge. ICANN (1) 2012: 482-490

2014: M. Franca, G. Zaverucha, A d'Avila Garcez, Fast relational learning using bottom clause propositionalization with artificial neural networks. Machine Learning 94(1): 81-104 (CILP++)

Neural-Symbolic Cycle



CILP++ system (Connectionist Inductive Logic Programming): download from http://sourceforge.net/projects/cilppp/

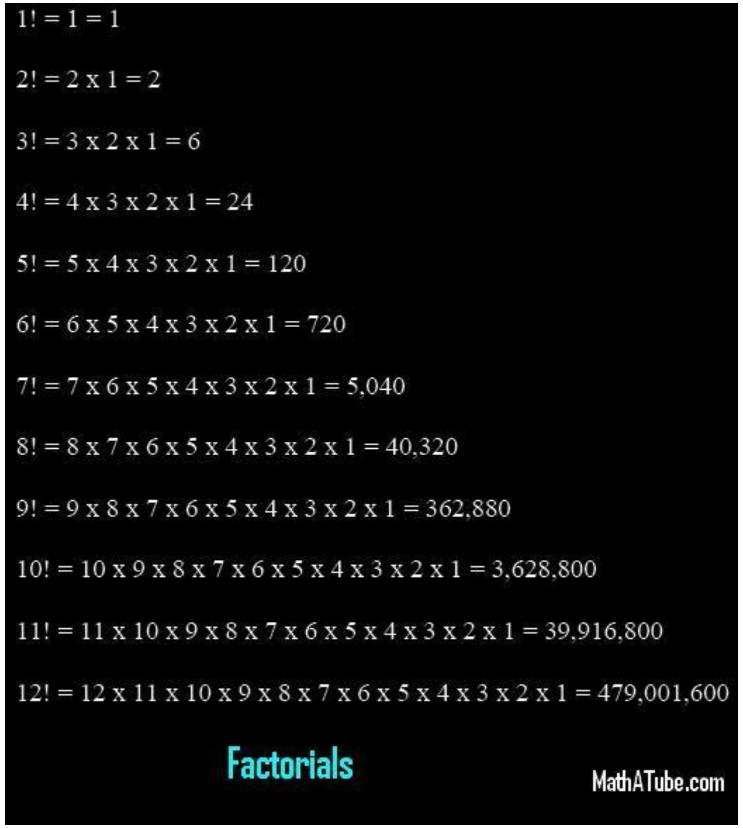
Goal-directed Reasoning and Learning

```
factorial(0,1).

factorial(N,F) :- N>0, N1 is N-1, factorial(N1,F1), F is N * F1.


?- factorial(3,W).

W=6
```



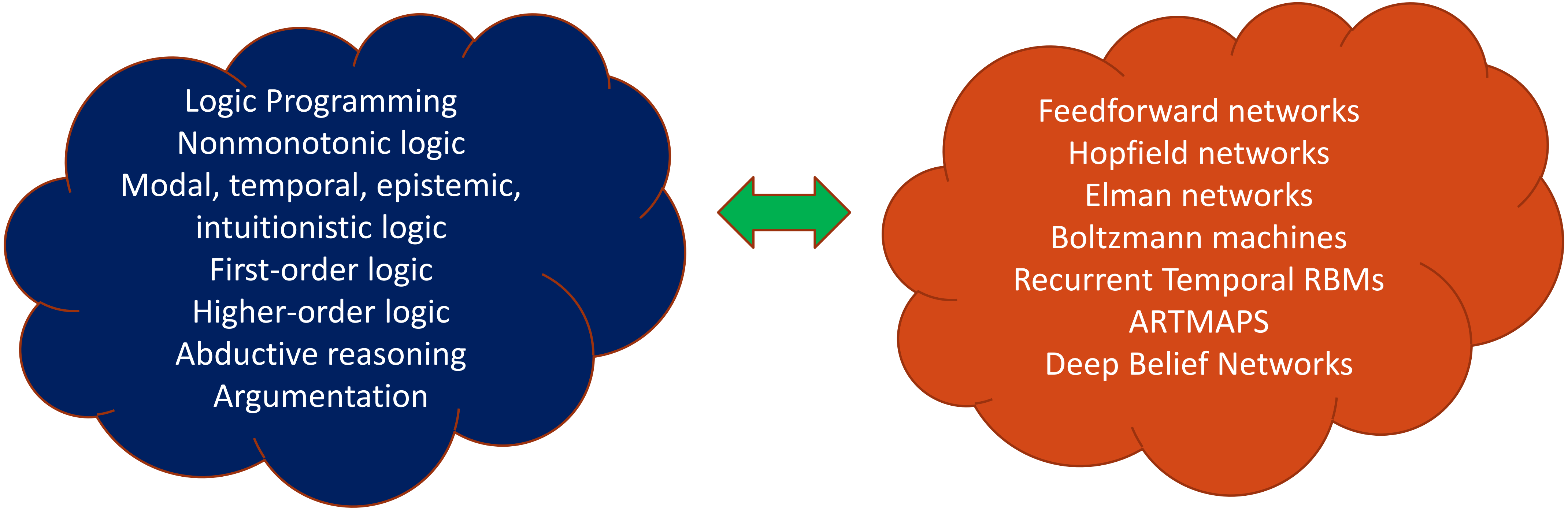
Inductive Logic Programming in a Neural Network (CILP++)

Teleo-reactive programs (N Nilsson) for cognitive robotics
(The order of the rules is important; higher rules have priority)

Tightly coupled and loosely coupled neural-symbolic systems

Neural-Symbolic Systems

Neural-symbolic methodology: translation algorithms to and from symbolic and connectionist models



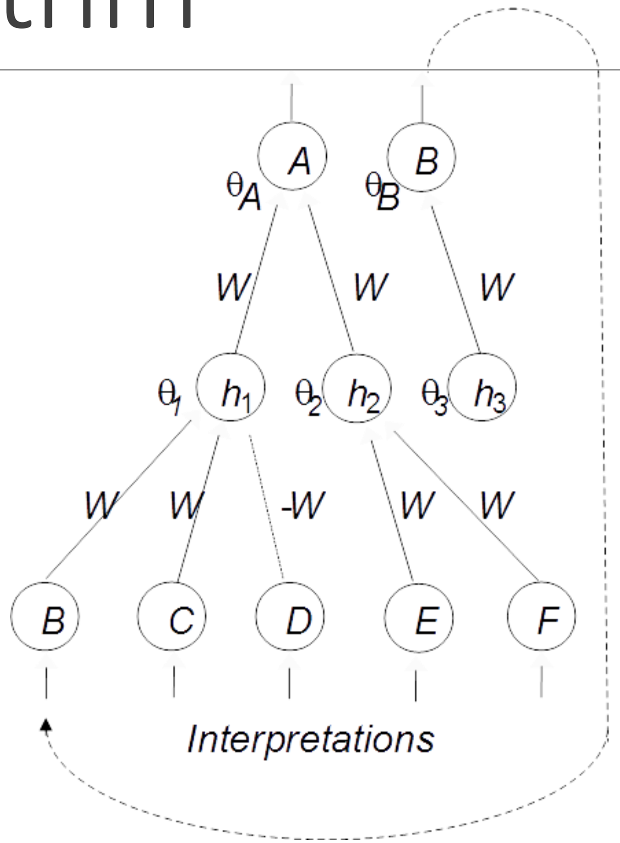
In search of robustness and explanations; “combining the logical nature of reasoning and the statistical nature of learning”, L. Valiant

CILP Translation Algorithm

r_1: $A \leftarrow B, C, \sim D;$

r_2 : $A \leftarrow E, F;$

r_3 : $B \leftarrow$



THEOREM 1: For any logic program P there exists a neural network N such that N computes P

Early application

Promoter = small DNA sequence at beginning of genes

Background Knowledge (14 rules):

```
Promoter IF Contact AND Conformation
Contact IF Minus10 AND Minus35
Minus35 IF @-36'ttgac'
Minus35 IF @-37'cttgac'
Conformation IF @-47'caa*tt*ac' AND @-22'g***t*c' AND @-8'gcgcc*cc'
```

10-fold cross-validation on set of 106 examples

CILP networks learn faster than backprop and KBANN, and perform slightly better than backprop and KBANN. We attribute this to the soundness of the CILP translation (i.e. the above theorem).

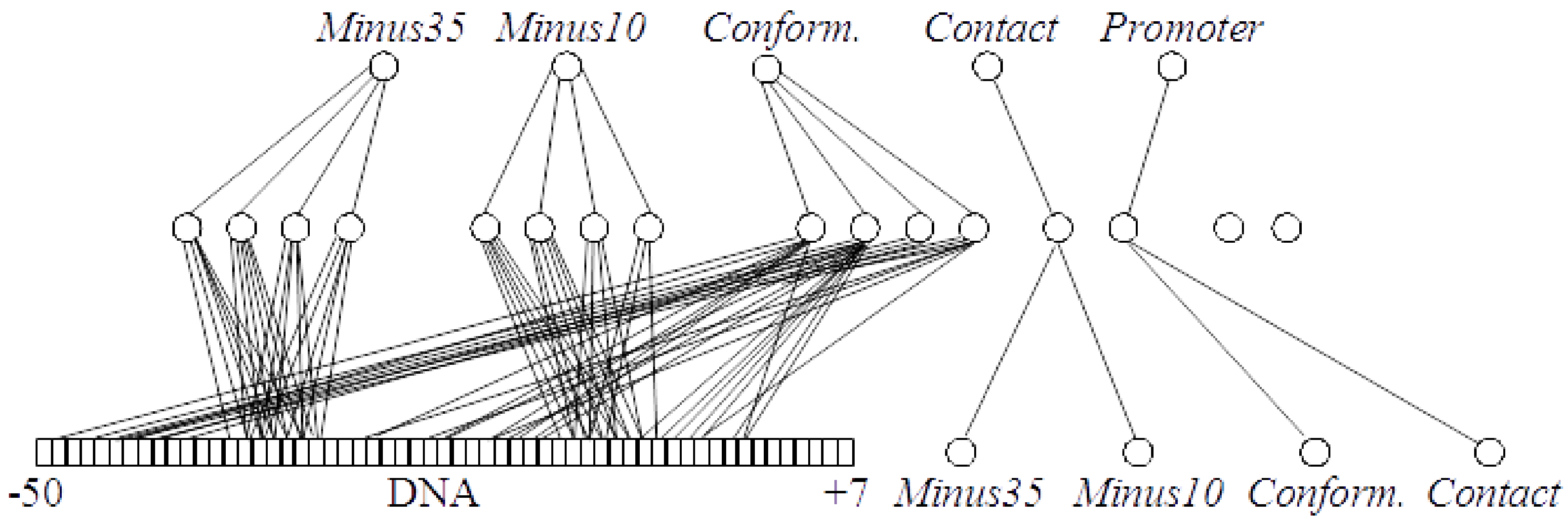
CILP Extraction Algorithm

Knowledge is extracted by querying the trained network;

A partial ordering helps guide the search, reducing complexity on the average case;

A proof of soundness guarantees that the rules approximate the network;

Rule simplification and visualization as state transition diagram can help experts validate results

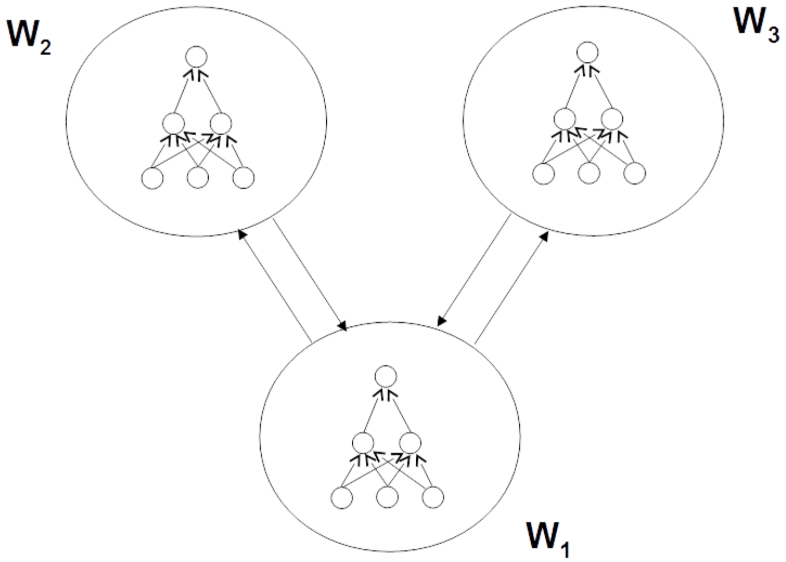


Extensions of CILP

Modal, Temporal, Epistemic, Intuitionistic, Preference Logic, Abductive Reasoning, Value-based Argumentation.

New applications include: normative reasoning (robocup), temporal logic learning (model checking), software model adaptation (business process evolution), training and assessment in simulators (driving test).

CILP network ensembles;
Represent and compute possible worlds;
Modularity for learning;
Modelling uncertainty through disjunctive information

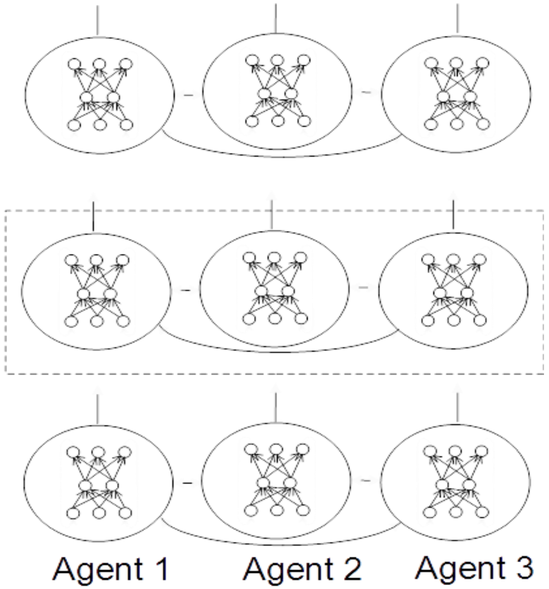


Connectionist Modal/Temporal Logic[1,2]

Use CILP network ensembles;

Propositional Modal Logic = decidable fragment of FOL with two variables

Full solution of Muddy Children puzzle



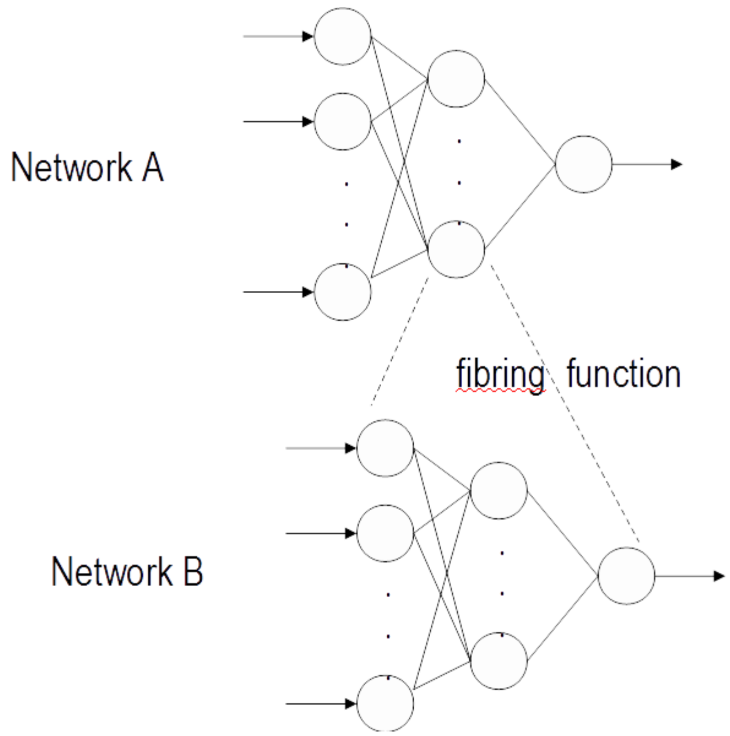
THEOREM 2: For any modal/temporal logic program *P* there exists an ensemble of neural networks *N* such that *N* computes *P*.

[1] Garcez, Lamb, Gabbay. Connectionist Modal Logic. Theoretical Computer Science, 371: 34-53, 2007.
[2] Garcez, Lamb. Connectionist Model for Epistemic and Temporal Reasoning. Neural Computation, 18:1711-1738, July 2006.

Fibring Neural Networks[3] (neuromodulation?)

A neuron that is a network!

Fibred networks approximate any polynomial function in an unbounded domain, as opposed to each of A, B,… which are universal approximators in compact domains.



[3] Garcez, Gabbay. Fibring Neural Networks. AAAI 2004, July 2004.

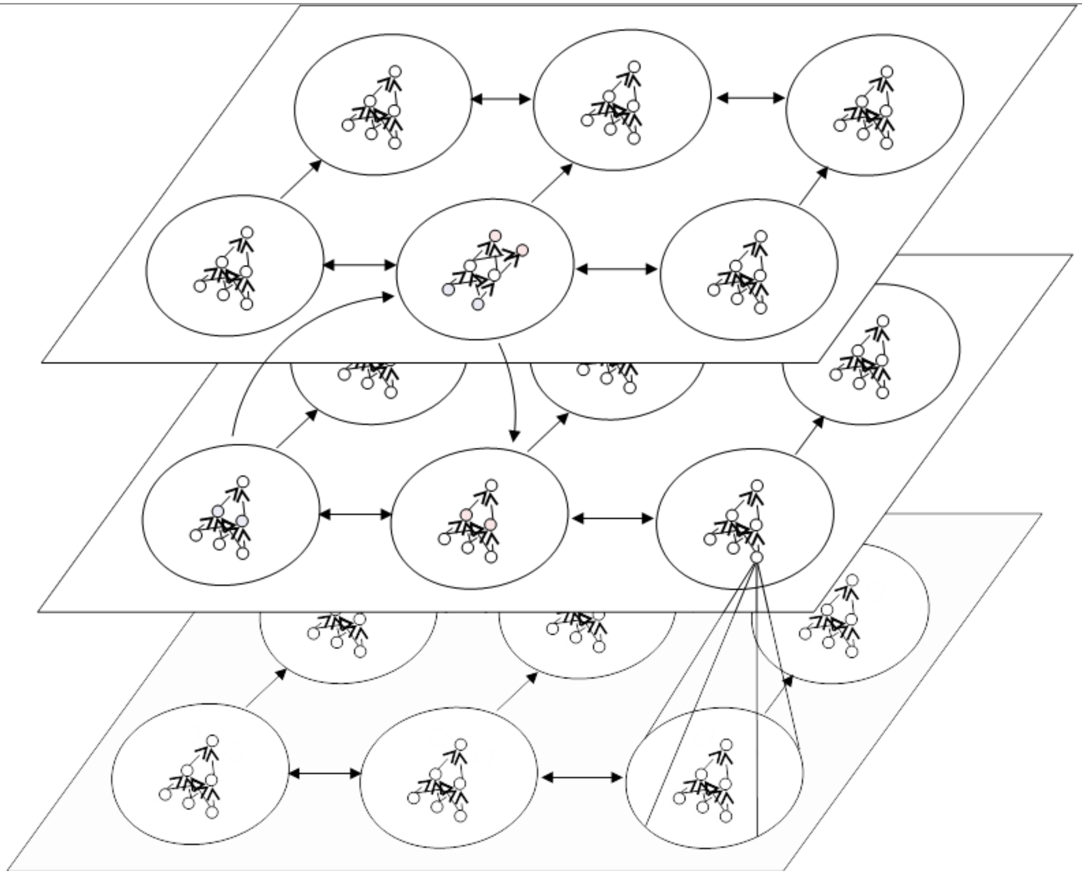
Fibred Network Ensembles[4]

Levels of abstraction

Modularity

Expressiveness

Relations and Specializations



[4] Garcez, Lamb, Gabbay. Neural-Symbolic Cognitive Reasoning. Springer, 2009.

Applications (1)

Training and Assessment in Simulators:

Learning new information from observation of experts and trainees at task execution and reasoning about this information online to provide feedback to the user

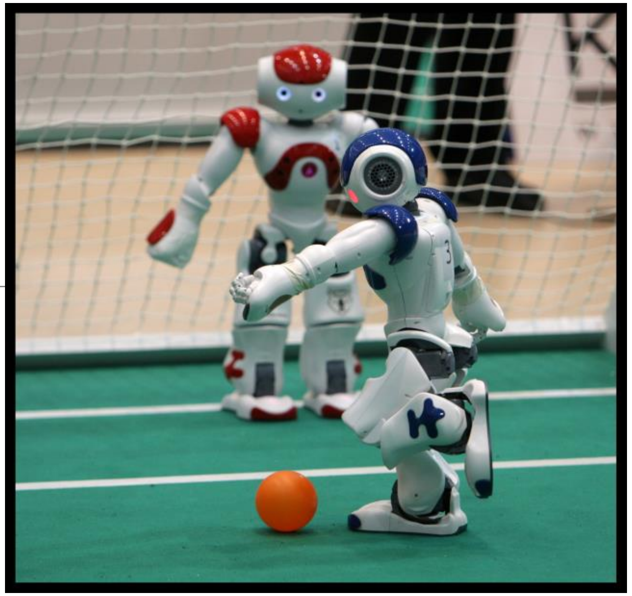
Uses knowledge insertion and extraction from RTRBMs

L. de Penning, A. S. d'Avila Garcez, L. C. Lamb and J. J. Meyer. A Neural-Symbolic Cognitive Agent for Online Learning and Reasoning. IJCAI'11, July 2011

Applications (2)

Learning Normative Rules of the RoboCup Competition

Software Model Verification and Adaptation



G. Boella, S. Colombo-Tosatto, A. S. d'Avila Garcez, V. Genovese, A. Perotti and L. van der Torre. Learning and Reasoning about Norms using Neural-Symbolic Systems. In Proc. AAMAS'12, Valencia, Spain, June 2012.

R. V. Borges, A. S. d'Avila Garcez and L. C. Lamb. Learning and Representing Temporal Knowledge in Recurrent Networks. IEEE Transactions on Neural Networks, 22(12):2409 - 2421, December 2011.

A. Perotti, A. S. d'Avila Garcez and Guido Boella, Neural Networks for Runtime Verification. The 2014 International Joint Conference on Neural Networks (IJCNN 2014), IEEE WCCI, Beijing, China, July 2014.

Relational Learning – CILP++

A neural-symbolic system that can solve ILP problems (by translating BK, mode declarations and ILP examples into vectors for training with backpropagation using CILP networks)

M. França, G. Zaverucha, A. S. d'Avila Garcez, Fast relational learning using bottom clause propositionalization with artificial neural networks. Machine Learning 94(1): 81-104 (2014)

Work in progress

Efficient First Order Logic Learning: encoding vs. propositionalisation?

- A number of attempts at first-order (and higher-order) logic learning via encoding: semantic approach (Hitzler et al), fibring, Topos (Osnabruck), association rules (SHRUTI), unification (Pinkas, etc.)
- Alternatively, relational learning via CILP++ (relevant for the analysis of complex networks: drug design in bioinformatics, link analysis in social networks, etc.)
- CILP++ (download it from sourceforge) competition: (probabilistic) ILP, MLNs, BLOG, StarAI (lifted inference)

Effective Knowledge Extraction from Very Large Networks: understanding multiple layers of representation, including semi-supervised learning and deep networks

- Visual intelligence applications: describing the actions in videos

Knowledge extraction from RBMs

[Pinkas 1995]: rule extraction from symmetric networks using penalty logic; proved equivalence between conjunctive normal form and energy functions

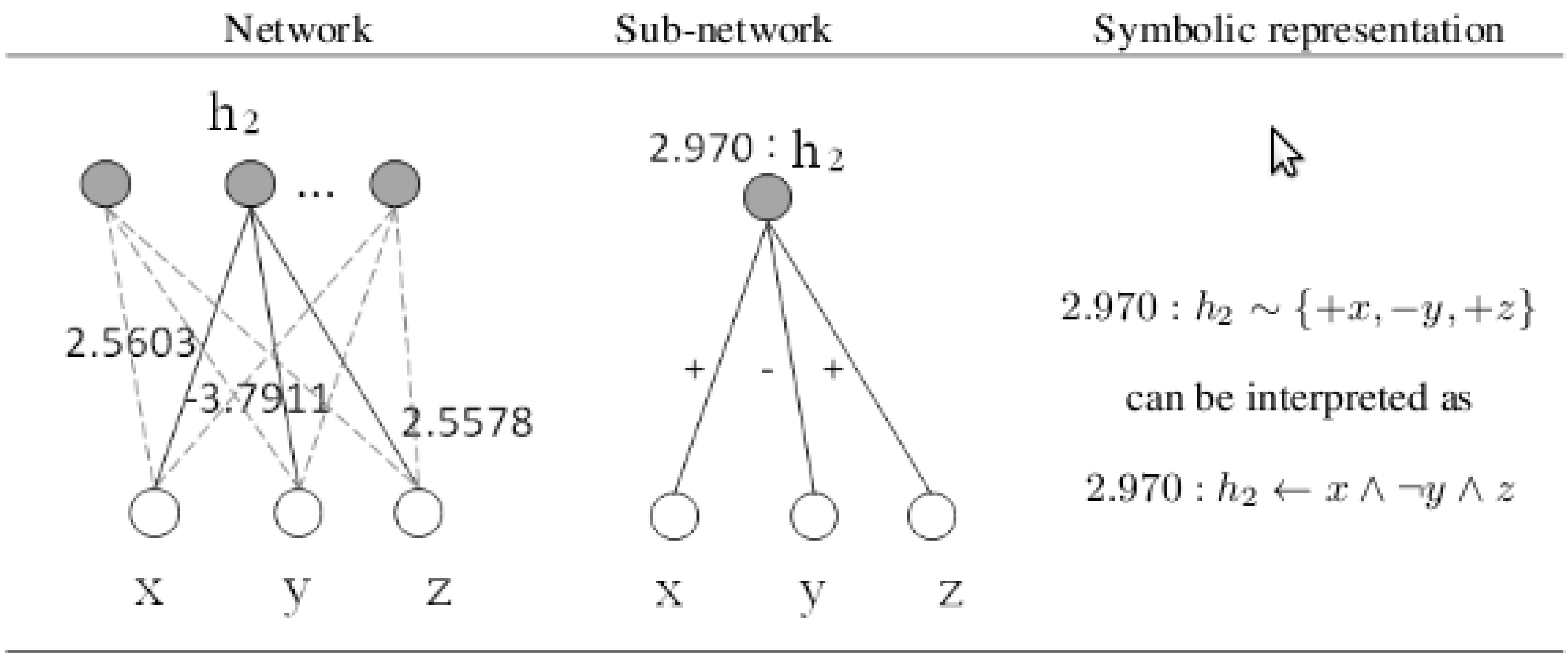
[Penning et al 2011]: extraction of temporal logic rules from RTRBMs using sampling; extracts rules of the form:

```
hypothesis(t) iff belief(1) Λ ... Λ belief(n) Λ hypothesis(t-1)
```

[Tran and Garcez, 2012, 2014]: rule extraction from DBNs using confidence-value similar to penalty logic but maintaining implicational form; extraction without sampling

Low-cost representation for RBMs…

Decompose the network into rules (pruning weights; averaging the absolute values of the remaining weights)



…guaranteed to minimise information loss

	TiCC	MNIST_{10K}	MNIST_{60K}	YALE face
RBM (J=500)	$94.851\% \pm 0.033$	97.198 ± 0.060	$98.553\% \pm 0.031$	$95.000\% \pm 2.833$
Rules	$94.711\% \pm 0.072$	97.240 ± 0.089	$98.530\% \pm 0.040$	$94.333\% \pm 3.865$
RBM (J=1000)	$94.928\% \pm 0.016$	$97.245\% \pm 0.031$	$98.680\% \pm 0.024$	$97.000\% \pm 2.919$
Rules	$94.729\% \pm 0.070$	$97.219\% \pm 0.056$	$98.562\% \pm 0.035$	$96.667\% \pm 1.757$

P Norvig, On Chomsky and the Two Cultures of Statistical Learning, Ninth Neural-Symbolic Learning and Reasoning Workshop NeSy13, IJCAI-13, Beijing, China, August 2013

Conclusion

To integrate the statistical nature of learning and the logical nature of reasoning

To develop effective computational systems for integrated reasoning and learning

To apply and evaluate such systems in complex new tasks…

Advertisements

Neural Relational Learning through Semi-Propositionalization of Bottom Clauses, M. Franca, A. d'Avila Garcez and G. Zaverucha (Poster at AAAI Spring KRR 2015)

Journal of Logic and Computation: learning and reasoning corner, A. d'Avila Garcez and L. Valiant (editors)

NeSy15 at IJCAI 2015, Buenos Aires, http://www.neural-symbolic.org/NeSy15/

NeSy 16 in New York (60 years of the Dartmouth conference)