Verification of Deep Neural Networks

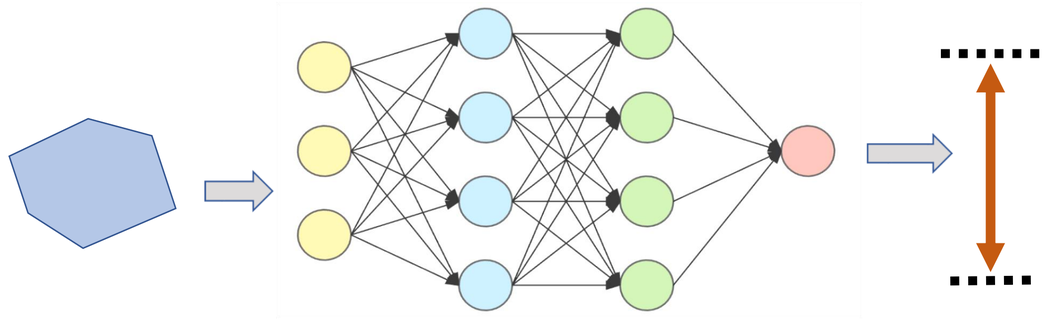
Souradeep Dutta, Sriram Sankaranarayanan
University of Colorado, Boulder

Background and Motivation

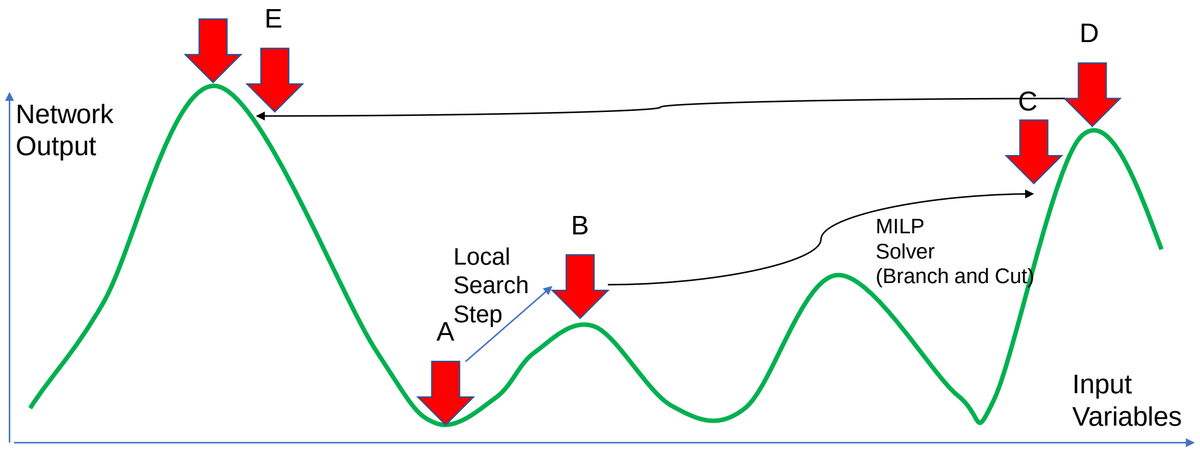
- Deep Neural Networks are used in high assurance systems such as autonomous drones, self-driving cars and closed loop medical devices.
- Traditional testing can barely scratch the set of behaviors for deep networks.
- How do we rigorously prove facts about the behavior of a deep network over a range of inputs?
- How do we train neural network controllers with baked in correctness guarantees?

Problem Statement

Input Set Propagation



Technique



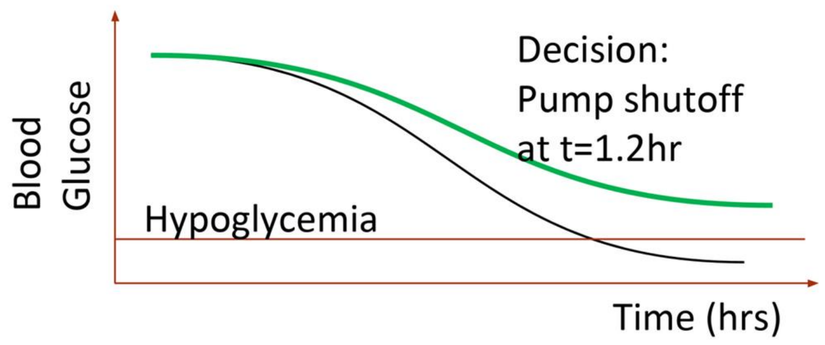
Sherlock Tool :
https://github.com/souradeep-111/sherlock

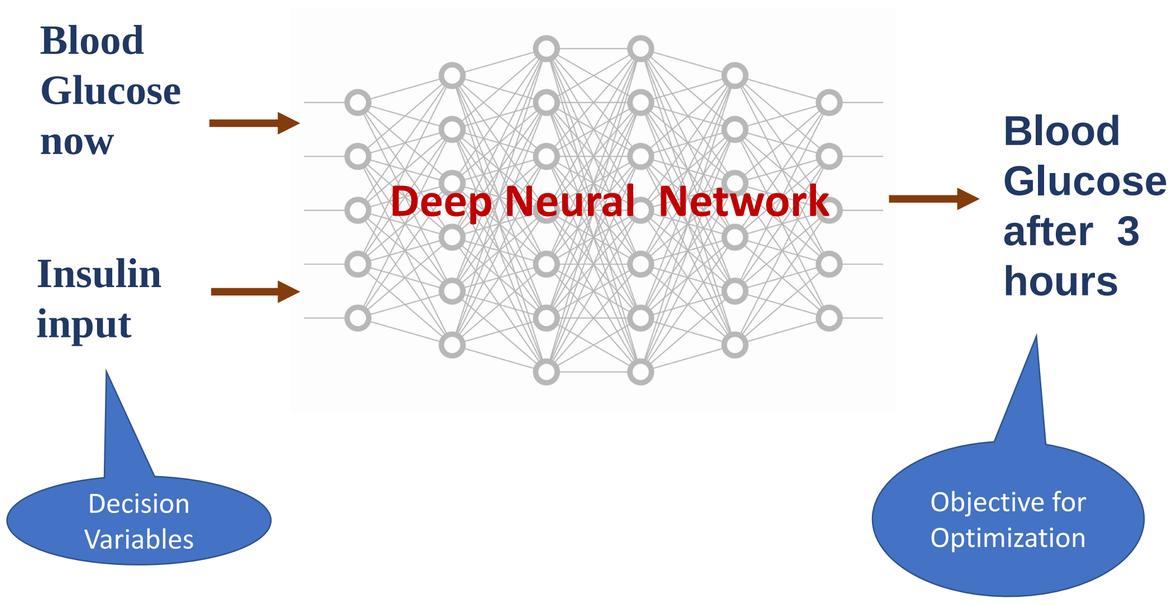
Results

#NEURONS	SHERLOCK TIME (seconds)	MONOLITHIC MILP TIME
100*	1	2.3
200*	2.2	3.6
425	4	6.1
478	6.5	40.8
500*	7.8	12.6
500*	1.5	0.5
731	3.7	3.9
762	227	281
778	18.3	71
1000*	232	232
1527	344	3312
2292	1457	13560
2340	3018	3626
3057	15000	219600
3822	259200	401700
6845	171	207

Applications

Deep Neural Networks as System Models





Future Research Challenges

- Deep Neural Networks used in autonomous systems are much larger than what can currently be handled by verification tools.
- Current approaches handle very simple specifications such as input/output ranges. Can we write specifications for more complex tasks such as image classification/segmentation?

Publications

- Formal Verification of neural network properties *(Dutta et al; NFM 2018)*.
- Synthesis of control strategies for neural network plant models (*Dutta et al; ADHS 2018)*.
- Control of blood glucose levels using neural network models for human patients (*Dutta et al; CMSB 2018)*.

Acknowledgments

Joint work with Drs. Ashish Tiwari and Susmit Jha. This work was funded in part by the US National Science Foundation (NSF) under award numbers CNS-1646556,