Formal Validation of Neural Networks as Timed Automata

서울시립대학교
UNIVERSITY OF SEOUL

Introduction

- We encode LIF networks into timed automata.
- We analyse some intrinsic properties of the proposed model.
- We encode in temporal logics all the behaviours.
- We prove limits of LIF model to hold for our model.

Introduction

Models	biophysically meaningful	tonic spiking	phasic spiking	tonic bursting	phasic bursting	mixed mode	spike frequency adaptation	class 1 excitable	class 2 excitable	spike latency	subthreshold oscillations	resonator	integrator	rebound spike	rebound burst	threshold variability	bistability	DAP	accommodation	inhibition-indiced spiking	inhibition-induced bursting	chaos	# of FLOPS
integrate-and-fire	-	+	-	-	-	-	-	+	-	-	-	-	+	-	-	-	-	-	-	-	-	-	5
integrate-and-fire with adapt.	-	+	-	-	-	-	+	+	-	-	-	-	+	-	-	-	-	+	-	-	-	-	10
integrate-and-fire-or-burst	-	+	+		+	-	+	+	-	-	-	-	+	+	+	-	+	+	-	-	-		13
resonate-and-fire	-	+	+	-	-	-	-	+	+	-	+	+	+	+	-	-	+	+	+	-	-	+	10
quadratic integrate-and-fire	-	+	-	-	-	-	-	+	-	+	-	-	+	-	-	+	+	-	-	-	-	-	7
Izhikevich (2003)	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	13
FitzHugh-Nagumo	-	+	+	-		-	-	+	-	+	+	+	-	+	-	+	+	-	+	+	-	-	72
Hindmarsh-Rose	-	+	+	+			+	+	+	+	+	+	+	+	+	+	+	+	+	+		+	120
Morris-Lecar	+	+	+	-		-	-	+	+	+	+	+	+	+		+	+	-	+	+	-	-	600
Wilson	-	+	+	+			+	+	+	+	+	+	+	+	+	+		+	+				180
Hodgkin-Huxley	+	+	+	+			+	+	+	+	+	+	+	+	+	+	+	+	+	+		+	1200

Preliminaries

Definition 1. A *timed automaton TA* is a tuple $(L, l^0, X, \Sigma, Arcs, Inv)$, where

- L is a set of locations with $l^0 \in L$ the initial one
- X is the set of clocks,
- Σ is a set of communication labels,
- $Arcs \subseteq L \times (G \cup \Sigma \cup U) \times L$ is a set of arcs between locations with a guard in G, a communication label in $\Sigma \cup \{\varepsilon\}$, and a set of variable upgrades (e.g., clock resets);
- $Inv : L \longrightarrow G$ assigns invariants to locations.

LIF Neuron

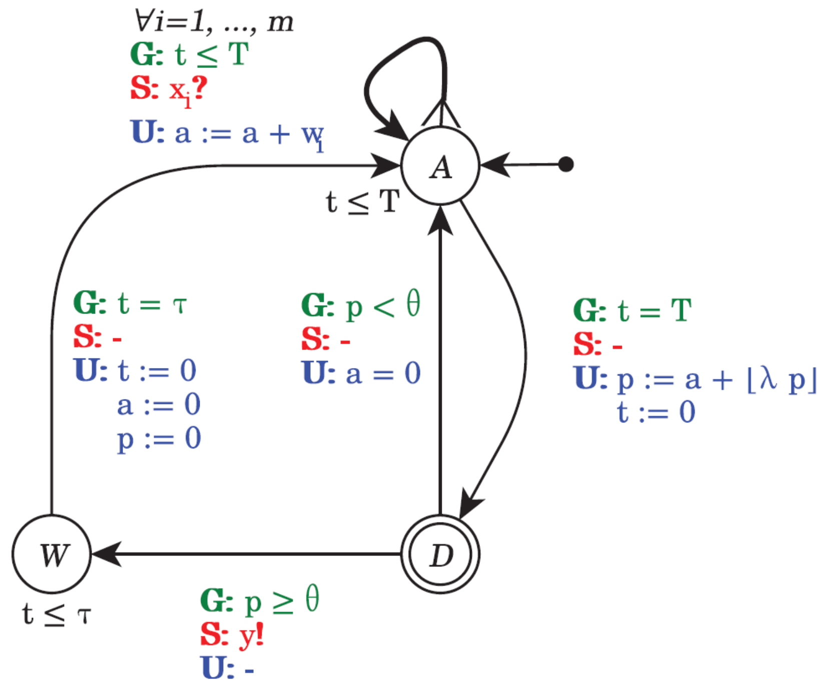

Figure 1: Neuron model.

where m is the number of input synapses, T is the accumulation period, and a is an integer variable storing the weighted sum of input spikes.

Properties

PROPERTY 4. *Let* $\mathcal{N} = (\theta, \tau, \lambda, p, y)$ *be a neuron and* a_{max} *be the maximum value received during each accumulation period. Then, if* $\theta \geq \frac{a_{max}}{1-\lambda}$*, the neuron is not able to fire.*

Properties

PROPERTY 5. *Let* $\mathcal{N} = (\theta, \tau, \lambda, p, y)$ *be a neuron. Then the time difference between successive firings cannot be lower than* $T + \tau$.

Properties

FACT 6. *Let* $\mathcal{N} = (\theta, \tau, \lambda, p, y)$ *be a neuron,* $a(t)$ *the sum of weighted inputs received during the current accumulation period, and* $p(t-1)$ *the neuron potential at the end of the previous accumulation period. If* $p(t-1) < \theta$ *and* $a(t) < 0$*, the neuron cannot fire at the end of the current accumulation period. Moreover, if* $p(t) \geq \theta$ *then* $a(t) > 0$*.*

Properties

Definition 7 (Inter-emission memory). Let $\mathcal{N}$ be a neuron, $Z_{\mathcal{N}}$ its reset times set, and I an input sequence. Then $\mathcal{N}$ has inter-emission *memory* if and only if there exist two different $t, t' \in Z_{\mathcal{N}}$ such that the output sequences produced by $\mathcal{N}$ as a response to I starting from t and t' are different.

PROPERTY 8. *Neurons have not inter-emission memory.*

VALIDATION OF THE MODEL

Tonic Spiking. *Tonic spiking* is the behaviour of a neuron producing a periodic output sequence as a response to a persistent excitatory constant input sequence.

Property 9 (Tonic spiking). *Let* $\mathcal{N} = (\theta, \tau, \lambda, p, y)$ *be a neuron having only one ingoing excitatory synapse of weight* w *and let* $\mathcal{I}$ *be the input source connected to* $\mathcal{N}$ *producing a persistent input sequence. Then* $\mathcal{N}$ *produces a periodic output sequence.*

VALIDATION OF THE MODEL

Integrator. *Integrator* is the behaviour of a neuron producing an output spike whenever it receives *at least* a specific number of spikes from its input sources in the same accumulation period.

PROPERTY 10 (INTEGRATOR). *Let $\mathcal{N} = (\theta, \tau, \lambda, p, y)$ be a neuron having m synapses with maximum excitatory weight R and a threshold $n \leq m$. Then the neuron emits if it receives a spike from at least n input sources during the same accumulation period.*

VALIDATION OF THE MODEL

Excitability. *Excitability* is the behaviour of a neuron emitting sequences having a *decreasing* inter-firing period, i.e., an increasing output frequency, when stimulated by an *increasing* number of excitatory inputs.

PROPERTY 11 (EXCITABILITY). *Let* $\mathcal{N} = (\theta, \tau, \lambda, p, y)$ *be a neuron having* m *excitatory synapses. Then the inter-spike period decreases as the sum of weighted input spikes increases.*

VALIDATION OF THE MODEL

Phasic Spiking. *Phasic spiking* is the behaviour of a neuron producing a *single* output spike when receiving a persistent and excitatory input sequence and then remaining quiescent for the rest of it. Such a behaviour depends on the neuron to have inter-emission memory.

PROPERTY 12. *Neurons cannot reproduce the* phasic spiking *behaviour.*

VALIDATION OF THE MODEL

Bursting. A *burst* is a finite sequence of *high frequency* spikes. More formally:

Definition 13. A spike output sequence is a *burst* if it is composed by spikes having an occurrence rate greater than $1/\tau$, with τ being the refractory period of the neuron.

PROPERTY 14. *Neurons cannot produce bursts.*

PROOF. A neuron N cannot emit spikes having a rate greater than $1/(T + \tau)$, as stated by Property 5, so it cannot produce bursts.

□

VALIDATION OF THE MODEL

Tonic Bursting is the behaviour of a neuron producing a burst sequence as a response to a persistent and excitatory input sequence. *Phasic Bursting* is the behaviour of a neuron producing a burst as a consequence of a persistent excitatory input sequence and then remaining quiescent. Obviously the preceding behaviours require the ability of producing bursts.

Bursting-then-Spiking is the behaviour of a neuron producing a burst as response to a persistent excitatory input sequence and then producing a periodic output sequence. Such a behaviour, similarly to Phasic and Tonic Bursting, depends on the neuron ability of producing bursts. Moreover it requires inter-emission memory, in order to detect the beginning of a persistent sequence.

PROPERTY 15. *Neurons cannot reproduce the* Tonic Bursting, Phasic Bursting *and* Bursting-then-Spiking *behaviour.*

VALIDATION OF THE MODEL

Spike Frequency Adaptation. *Spike Frequency Adaptation* is the behaviour of a neuron producing a decreasing-frequency output sequence as a response to a persistent excitatory input sequence. In other words, the inter-emission time difference increases as the time elapses. This behaviour requires the neuron to have inter-emission memory as it should be able to keep track of the time elapsed since the beginning of the input sequence.

PROPERTY 16. *Neurons cannot reproduce the* Spike Frequency Adaptation *behaviour.*

VALIDATION OF THE MODEL

Spike Latency. *Spike Latency* is the behaviour of a neuron firing delayed spikes, with respect to the instant when its potential reached or overcame the threshold. Such a delay is proportional to the strength of the signal which leads it to emission, i.e., the sum of weighed inputs received during the accumulation period preceding the emission. This behaviour requires the neuron to be able to postpone its output.

PROPERTY 17. *Neurons cannot reproduce the* Spike Latency *behaviour.*

VALIDATION OF THE MODEL

Threshold Variability. *Threshold variability* is the behaviour of a neuron allowing its threshold to vary according to the strength of its inputs. More precisely, an excitatory input will rise the threshold while an inhibitory input will decrease it. As a consequence, excitatory inputs may more easily lead the neuron to fire when occurring after an inhibitory input.

PROPERTY 18. *Neurons cannot reproduce the* Threshold Variability *behaviour.*

VALIDATION OF THE MODEL

Bistability. *Bistability* is the behaviour of a neuron alternating between two operation modes: *periodic emission* and *quiescence*. Upon reception of a single excitatory spike, it emits a periodic output sequence and switches to a quiescent mode (no emission) as soon as it received another spike. Such a behaviour requires the neuron to (i) be able to produce a periodic output sequence, even if no excitatory spike is received, (ii) be able to remain silent when no spike is received, and (iii) be able to switch between the two operation modes upon reception of an excitatory spike.

PROPERTY 19. *Neurons cannot reproduce the* Bistability *behaviour.*

VALIDATION OF THE MODEL

Inhibition-induced activities. *Inhibition-induced Spiking* is the behaviour of a neuron producing a spike output sequence as a response to a persistent *inhibitory* input sequence. We thus require the neuron to be able to emit as a consequence of some inhibitory input spikes.

PROPERTY 20. *Neurons cannot reproduce the* Inhibition-induced Spiking *behavior.*

PROOF. Follows from Fact 6. □

An easy extension to our automata is to consider *the absolute value* of all inputs instead of their signed values.

Rebound activities. *Rebound Spike* is the behaviour of a neuron producing an output spike after it received an inhibitory input. Similarly to Inhibition-induced activities, this behaviour requires the neuron to emit as a consequence of an inhibitory input spike.

PROPERTY 21. *Neurons cannot reproduce the* Rebound Spiking *behaviour.*

PROOF. Follows from Fact 6. □

Conclusion

- Our work is the first Timed Automata analysis of spiking neural network.