

Handling of Time

ANN fk
Örjan Ekeberg
Feed-Forward Networks
Recurrent Networks
Training of Recurrent Networks

Feed-Forward Networks

Tapped Delay Line

Time Lagged Feed-Forward Networks

Recurrent Networks

Different Architectures

Finite-State Machines

Training of Recurrent Networks

Back-Propagation Through Time

Teacher Forcing

ANN fk
Örjan Ekeberg
Feed-Forward Networks
Recurrent Networks
Training of Recurrent Networks

Feed-Forward Networks

Tapped Delay Line

Time Lagged Feed-Forward Networks

Recurrent Networks

Different Architectures

Finite-State Machines

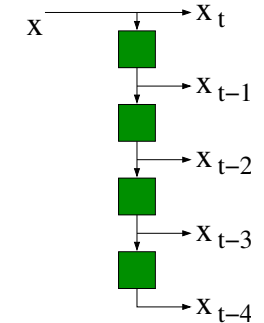
Training of Recurrent Networks

Back-Propagation Through Time

Teacher Forcing

ANN fk
Örjan Ekeberg
Feed-Forward Networks
Recurrent Networks
Training of Recurrent Networks

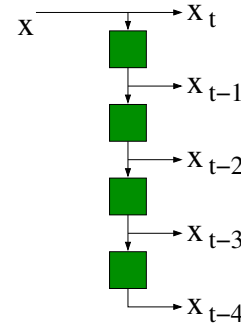
- ▶ Temporal signals are tricky
- ▶ Transform temporal signals to spatial
- ▶ Tapped Delay Line



ANN fk
Örjan Ekeberg
Feed-Forward Networks
Recurrent Networks
Training of Recurrent Networks

We must choose

- Memory depth
Number of steps saved
- Resolution
Size of time-steps

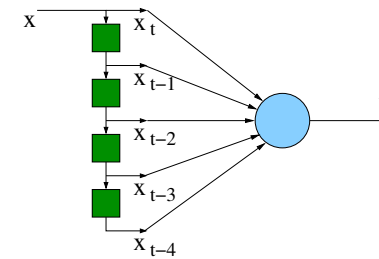


ANN fk

Örjan Ekeberg

Feed-Forward
Networks
Tapped Delay Line
Time Lagged
Feed-Forward
Networks
Recurrent
Networks
Training of
Recurrent
Networks

Simplest form
"Neuronal filter"

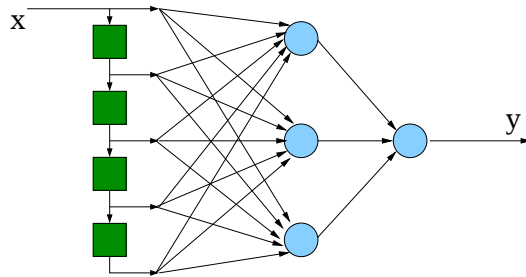


ANN fk

Örjan Ekeberg

Feed-Forward
Networks
Tapped Delay Line
Time Lagged
Feed-Forward
Networks
Recurrent
Networks
Training of
Recurrent
Networks

Multi-Layer Networks



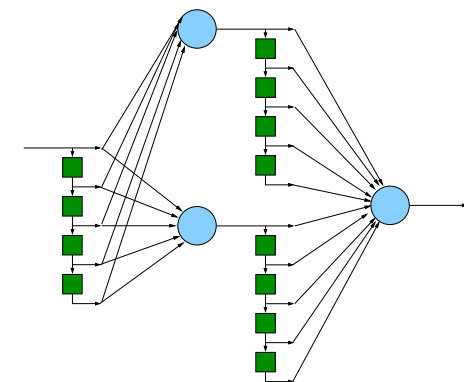
ANN fk

Örjan Ekeberg

Feed-Forward
Networks
Tapped Delay Line
Time Lagged
Feed-Forward
Networks
Recurrent
Networks
Training of
Recurrent
Networks

Delay of **all** signals

- More general
- Harder to train



ANN fk

Örjan Ekeberg

Feed-Forward
Networks
Tapped Delay Line
Time Lagged
Feed-Forward
Networks
Recurrent
Networks
Training of
Recurrent
Networks

- ▶ How can we train a network with delays?
- ▶ Ordinary Back-Prop?
- ▶ Works when only the input is delayed
- ▶ Generalizations are needed when general delays are included
- ▶ The error signal must be matched against **old** activity values

ANN fk
Örjan Ekeberg

- Feed-Forward Networks
 - Tapped Delay Line
 - Time Lagged Feed-Forward Networks
- Recurrent Networks
- Training of Recurrent Networks

Feed-Forward Networks
Tapped Delay Line
Time Lagged Feed-Forward Networks

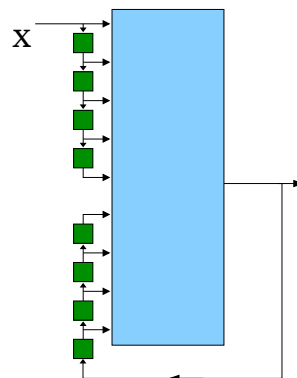
Recurrent Networks
Different Architectures
Finite-State Machines

Training of Recurrent Networks
Back-Propagation Through Time
Teacher Forcing

ANN fk
Örjan Ekeberg

- Feed-Forward Networks
- Recurrent Networks
 - Different Architectures
 - Finite-State Machines
- Training of Recurrent Networks

Feedback of the output

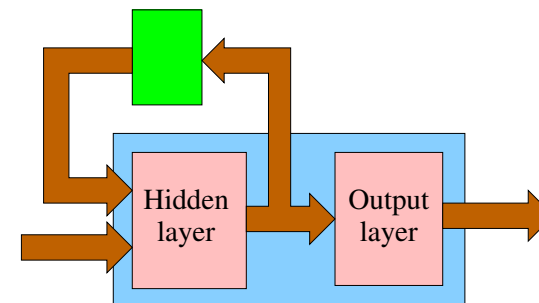


NARX — Nonlinear AutoRegressive with eXogenous inputs

ANN fk
Örjan Ekeberg

- Feed-Forward Networks
- Recurrent Networks
 - Different Architectures
 - Finite-State Machines
- Training of Recurrent Networks

Feedback from the hidden layer in next time-step

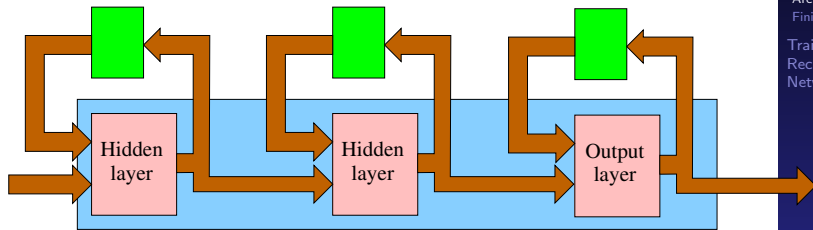


Context Units

ANN fk
Örjan Ekeberg

- Feed-Forward Networks
- Recurrent Networks
 - Different Architectures
 - Finite-State Machines
- Training of Recurrent Networks

Multi-Layered Recurrent Networks



Recurrent Multi-Layer Perceptron

What can a recurrent network do?

- ▶ Finite-State Machine
- ▶ State corresponds to activity in the context nodes
- ▶ Combination of state and input produces next state

Second Order Networks

- ▶ Product of input signal and feedback
- ▶ Every pair: $\langle \text{in-signal} \times \text{feedback} \rangle$ has its own weight
- ▶ Many (specific) weights

Feed-Forward Networks

Tapped Delay Line

Time Lagged Feed-Forward Networks

Recurrent Networks

Different Architectures

Finite-State Machines

Training of Recurrent Networks

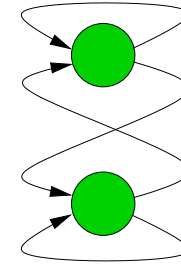
Back-Propagation Through Time

Teacher Forcing

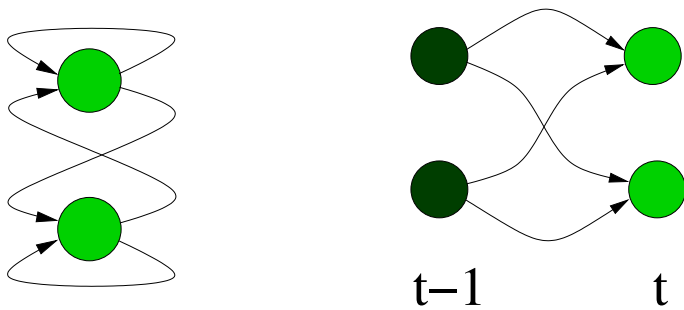
Is it possible to train a recurrent network to reproduce given sequences?

► Back-Propagation Through Time

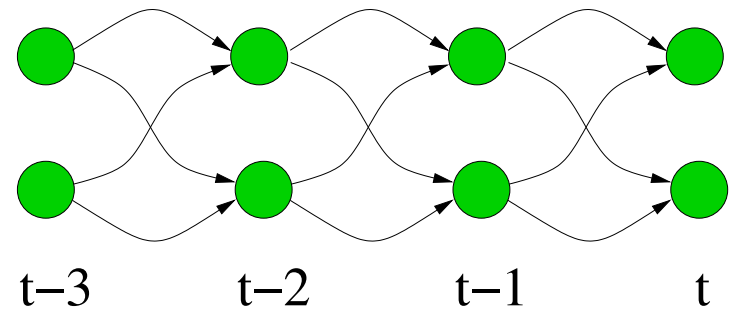
Recurrent network with delayed connections

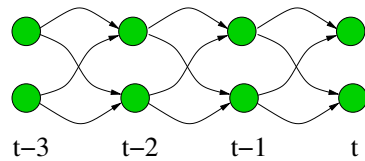


Current state depends (only) on the previous state

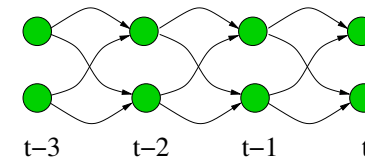


Unfolding of the History





- ▶ Back-Propagation Through Time
- ▶ Target values are spread out over different layers
- ▶ Add contributions from all goals
- ▶ The same weight occurs in multiple places!
- ▶ Add contribution from all places



- ▶ All old activity values must be stored
- ▶ Problematic when learning long sequences
- ▶ Truncation of the history

- ▶ Continuous Learning
- ▶ Corresponds to Incremental Learning
- ▶ Only works if the time-steps are small

Teacher Forcing

- ▶ Back-Prop uses activity produced by the net
- ▶ Why not use “true value” when available?
- ▶ Acceleration by using old goal values