MULTCONS Philosophy

Multilayer Temporal Constraint Satisfaction Network

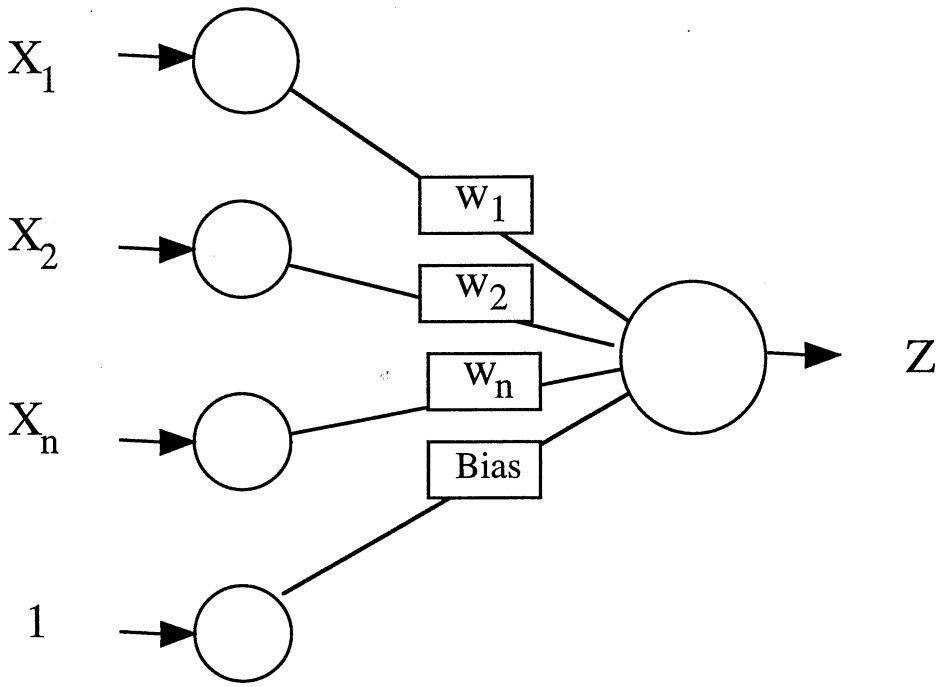
- Fundamental limits on how accurate phonetic "labeling" can be, based only on acoustic features - even "if" perfect labeler, we do not speak phonetically correct

- Requires context
  - Local acoustic context - time-varying features - recurrent network/Input preprocessing
  - Phoneme context
  - Word context
  - Phrase/Grammar context
  - Topic/Application context

- MULTCONS allows all levels of context to interact and compete before final decisions/suggestions are made

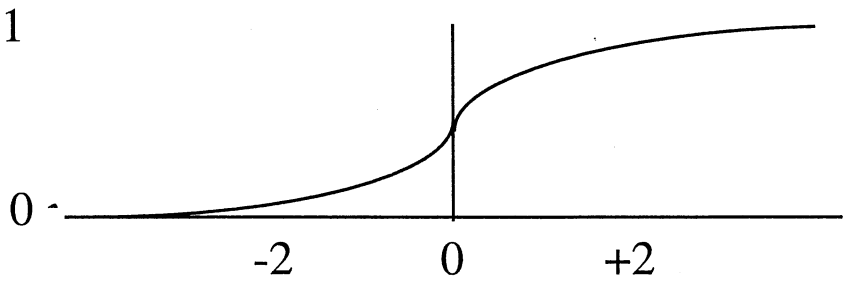
- Based on a Novel Neural Network Approach

Basic Neuron Node



$$net = \sum x_i w_i$$

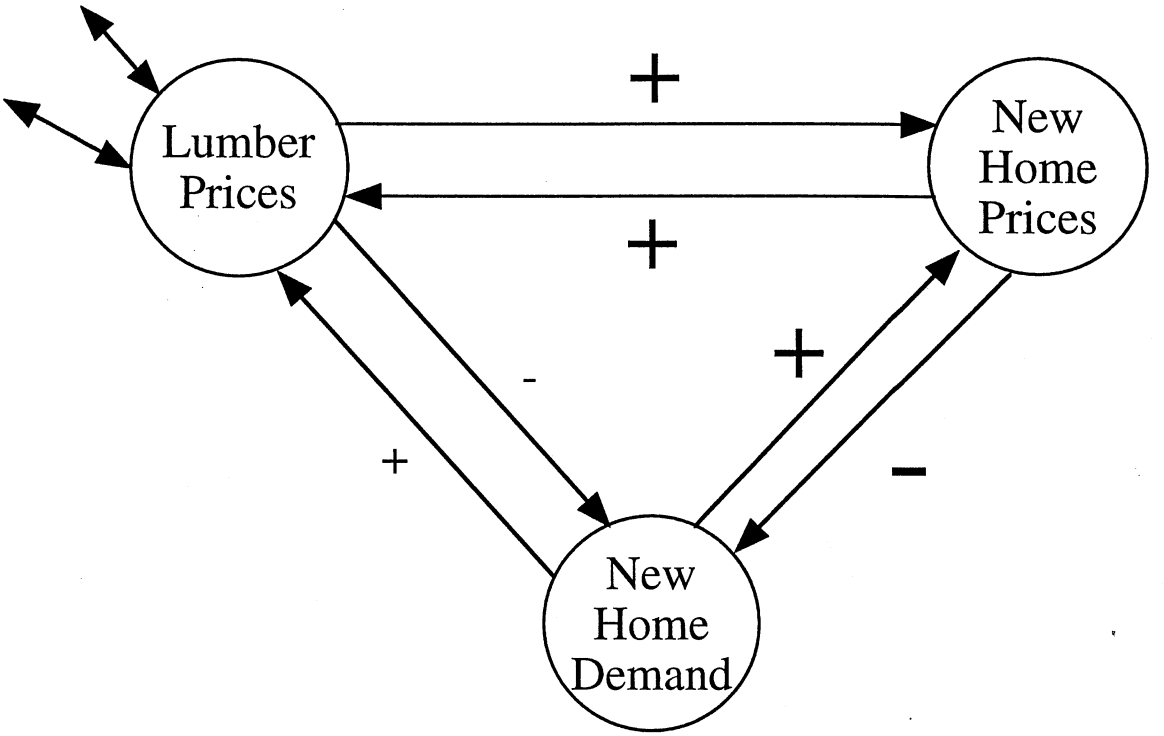
Node Activation = f(net)



Activation Function

Learns by Modifying Weights after seeing training data, then generalizes on novel data (i.e. medical diagnosis)

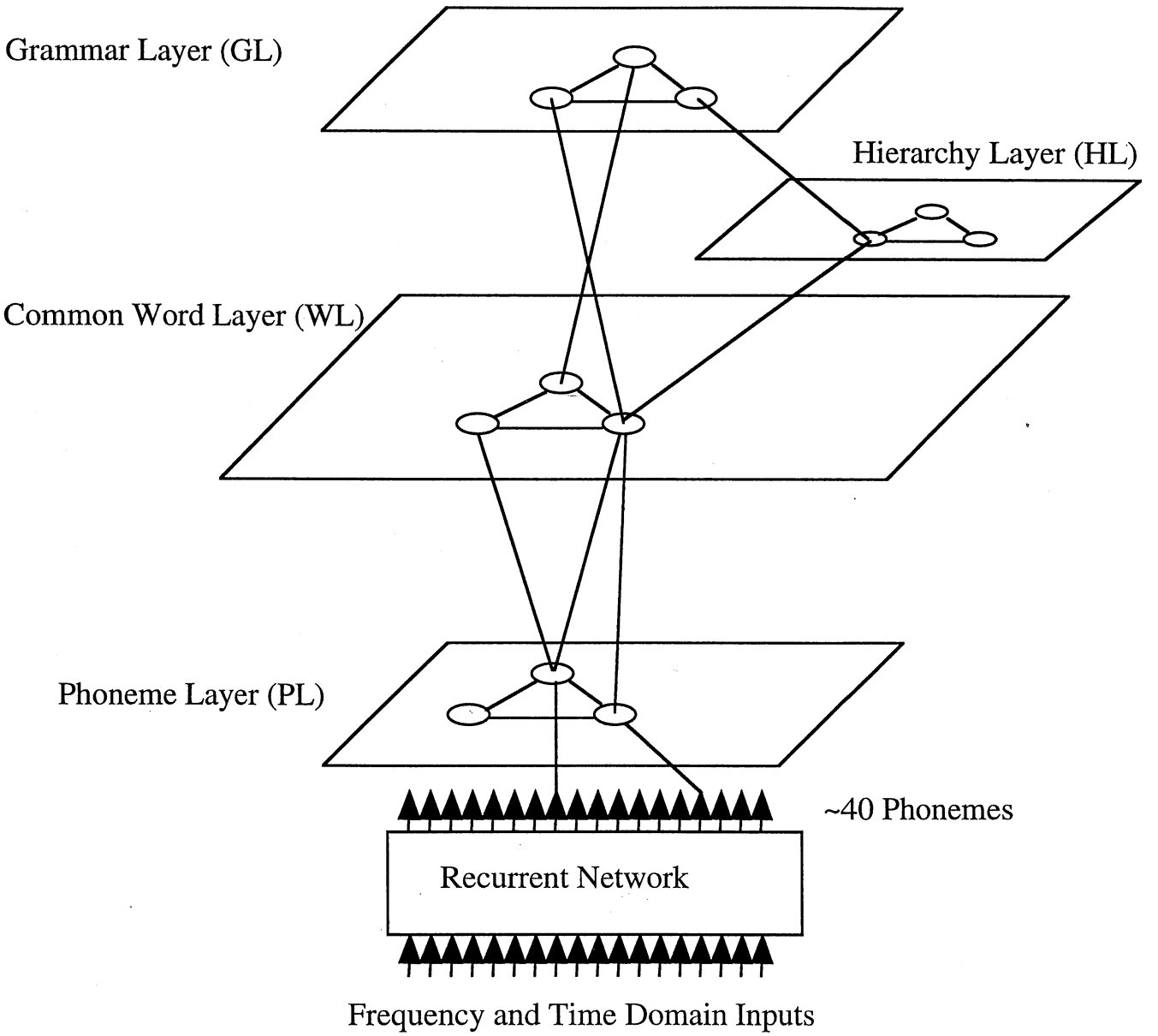
Constraint Satisfaction Network Example: Free Market Economics



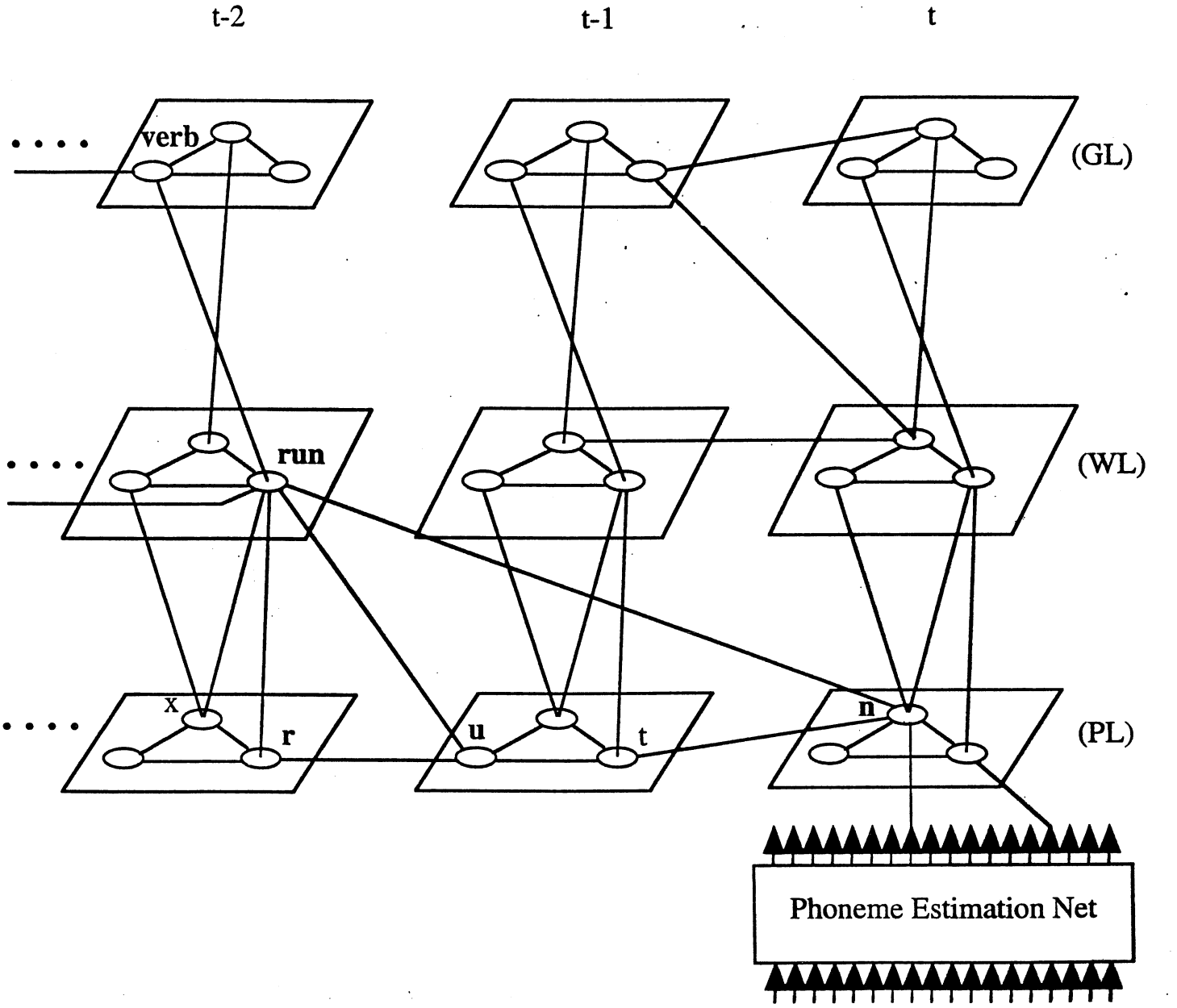
- Typically stable at an equilibrium state
- Assume that an outside adjustment is made (e.g. Spotted Owl)
- After a short amount of interaction, the network relaxes to a new stable state

MULTCONS

MUlti-layer Temporal CONstraint Satisfaction networks



MULTCONS Timing

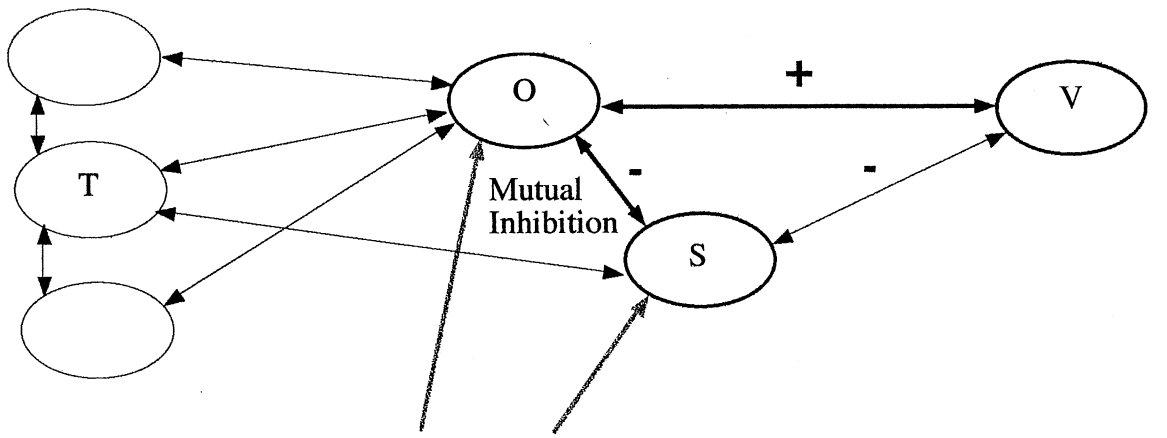


Phoneme Time Step Triggered by
Segmentation Estimation Algorithm

Phoneme Layer Interaction

Preceding Phonemes

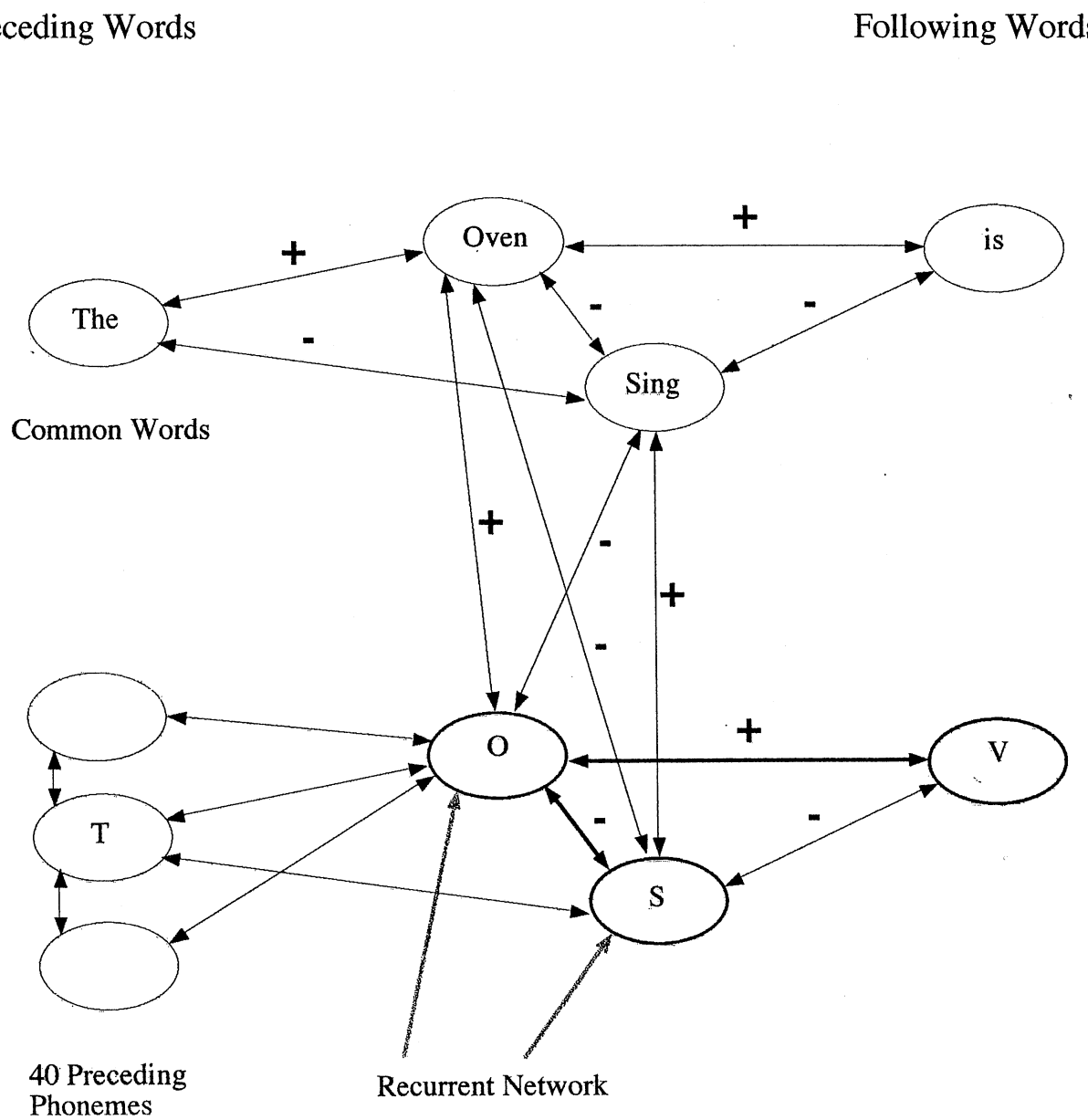
Following Phonemes



40 Preceding Phonemes

Recurrent Network

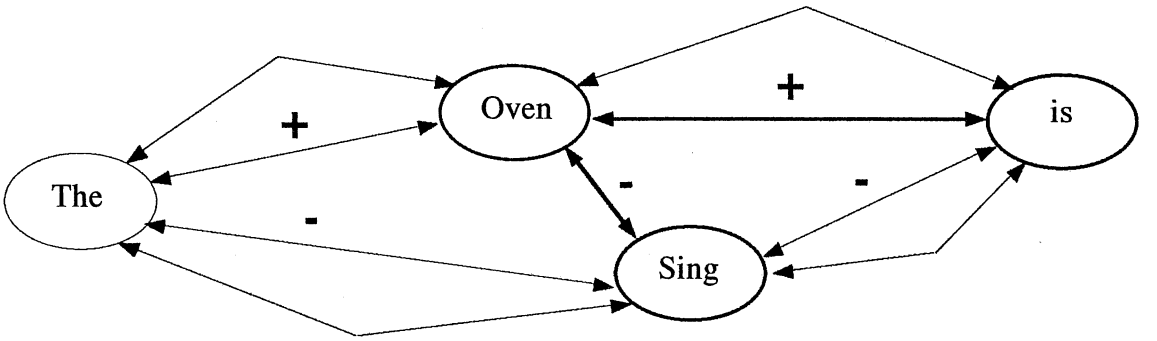
Phoneme Layer and Word Layer Interaction



Word Layer Sequential Excitation

Preceding Words

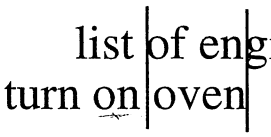
Following Words



Common Words

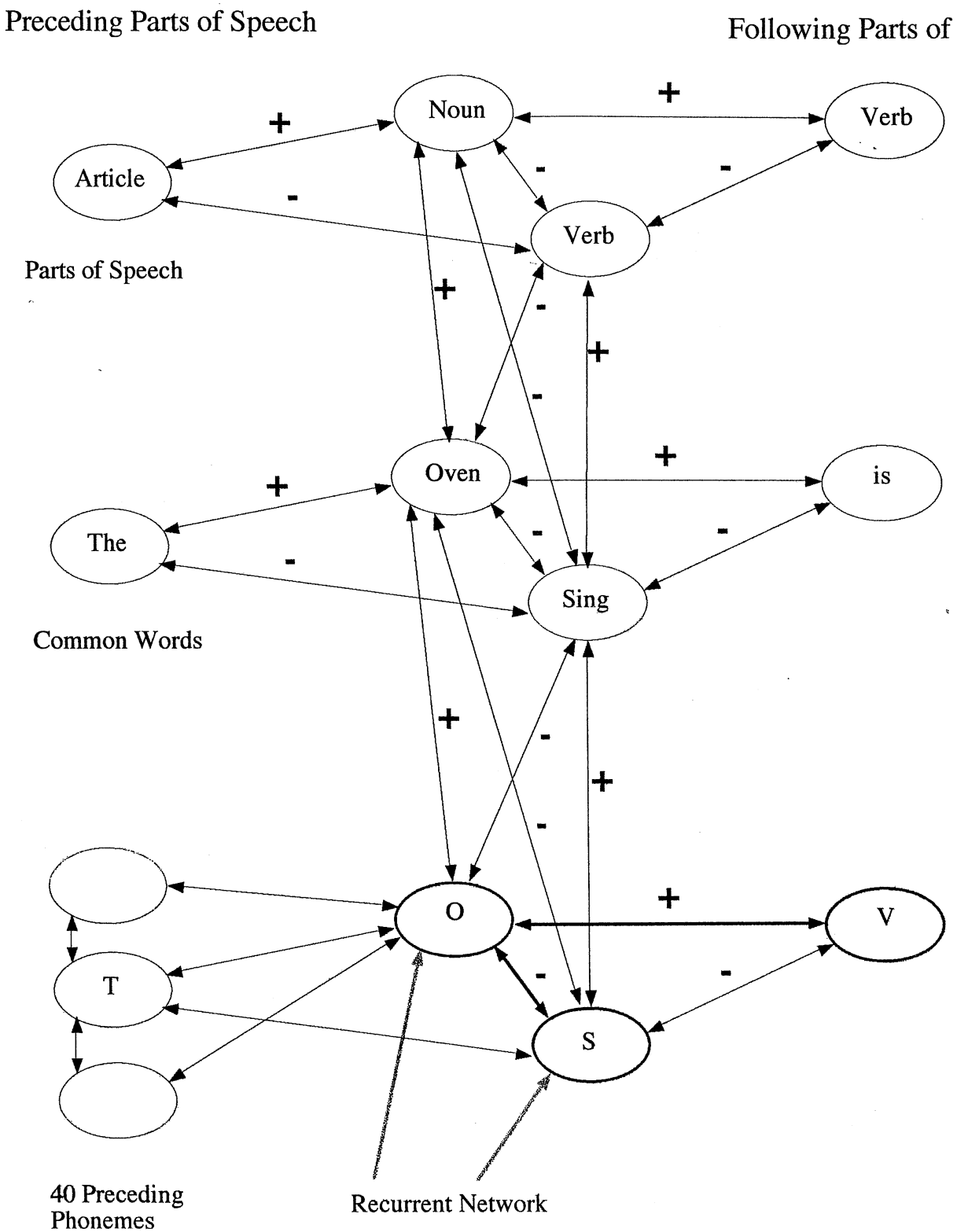
Grammar Layer Usage

Assume possible competing words "of" and "oven" at the word layer:

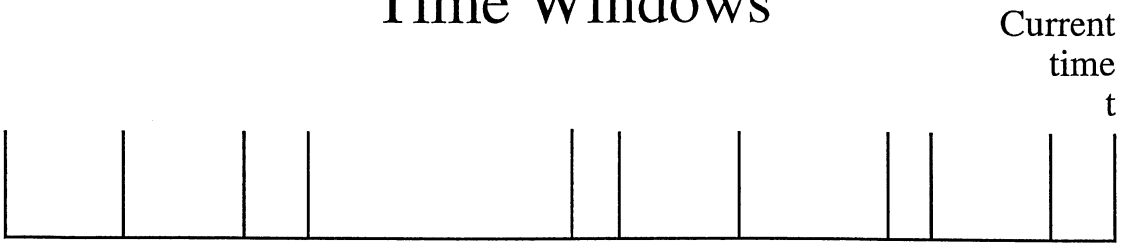


If the word "on" is active in the word layer then the corresponding "Preposition" node of the grammar layer would be active. Since nouns typically follow prepositions, the following "noun" node in the grammer layer would be active and would excite the "oven" node in the word layer. The "of" node of the word layer would be inhibited since it's corresponding preposition node would be inhibited by the preposition node corresponding to "on".

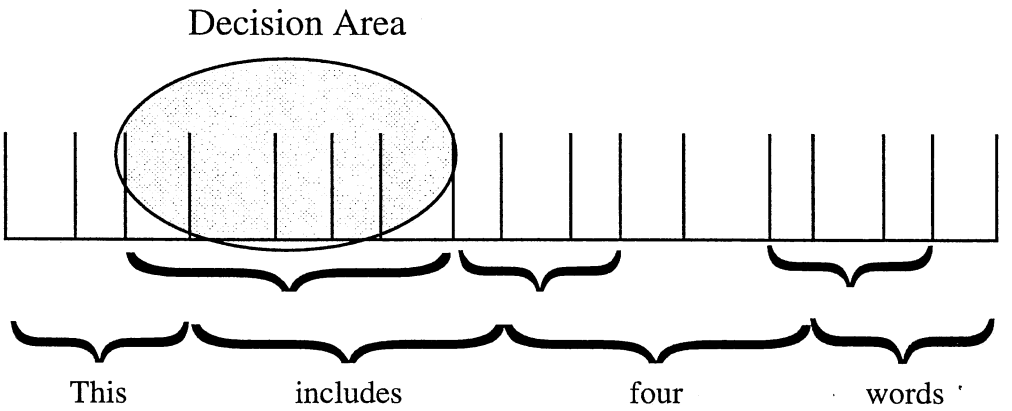
Addition of Grammar Layer



Time Windows

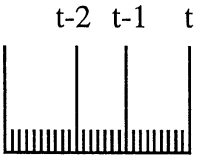


Sentence Layer - only competing word slots, slightly longer or the same as the word layer. Note that in this case final word decision lags about 3 words behind time t. The decision can be the single best word sequence or the top c candidates. Sentence and Word windows slide left on phoneme shifts and words continue to interact until they are completely out of the window.



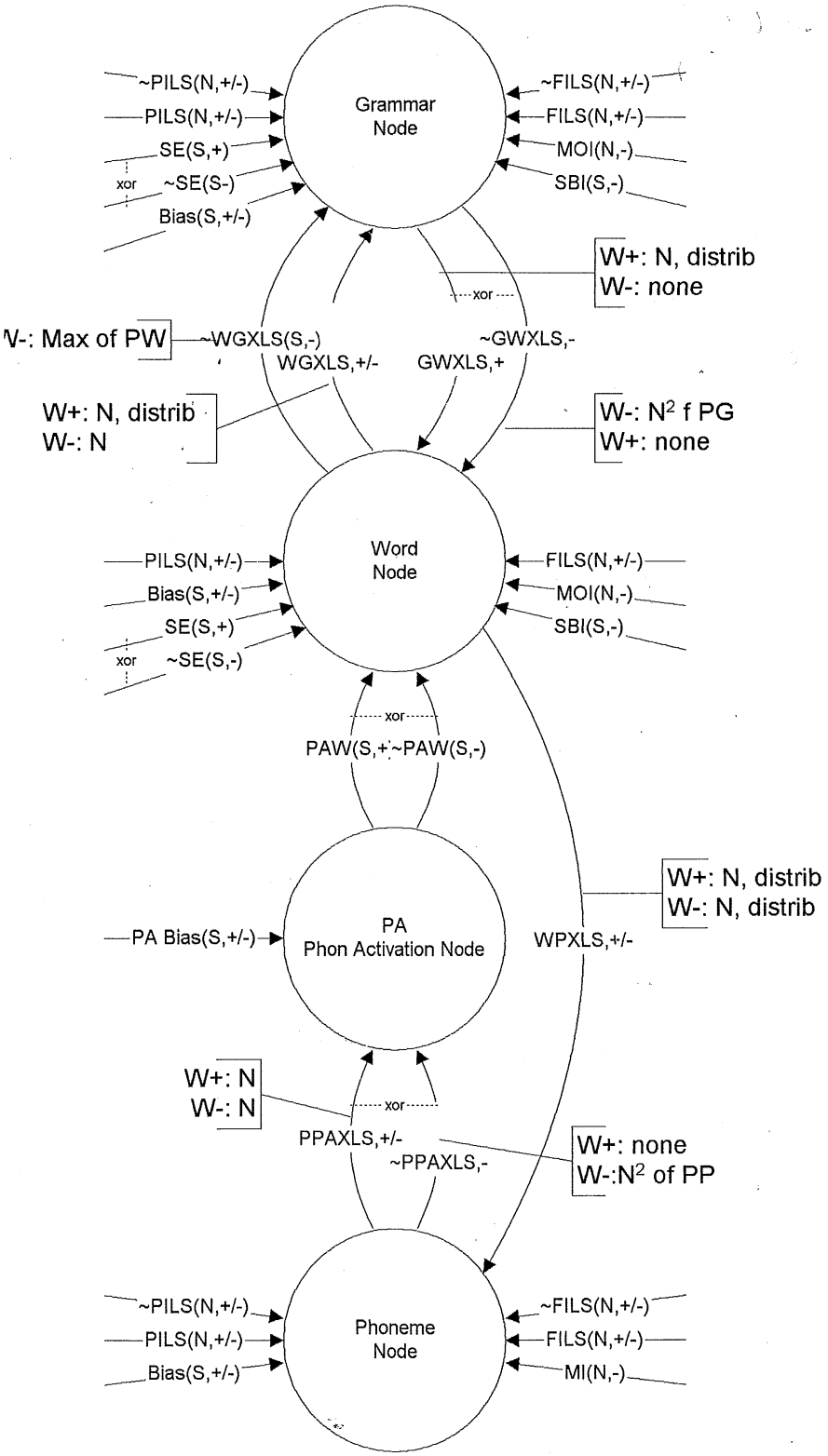
Word Layer - both phoneme and word slots, potential words overlap - length of w phoneme slots (for w=25, covers approximately 5 words). Note that contiguous streams of words are best candidates. This is supported automatically by the network interaction. Also note that the left half of the word window has no vertical interaction with the PL, but continues to have horizontal interaction within the WL and vertical interaction with the SL.

Phoneme Layer - k phoneme slots. k an odd value where $5 \leq k \leq 11$. Run M network time steps before accepting the next phoneme shift. Note that earlier phonemes will be gone before final word decisions are made.



Recurrent Net - Microbes and Phoneme time slot generation

Multcons Connections



FILS - Follows Intralayer Synapse. Weight based on the probability of Y given that X follows.
~FILS - Not Follows ILS. Weight based on the probability of Y given that X does not follow.
PILS - Preceeds Intralayer Synapse. Weight based on the probability of Y given that X preceeds.
~PILS - Not Preceeds ILS. Weight based on the probability of Y given that X does not preceed.
SE - Sequential Excitation. Weight based on a sequence of active nodes or properly bounded by silence.
SBI - Silence boundary inhibition. Inhibition due to a silent area found in the middle of a word.
MOI - Mutual Overlap Inhibition. Inhibitory weight between overlapping nodes based on activations and amount of overlap.
MI - Mutual Inhibition. Inhibitory weight based on excitation of other phonemes.
Bias - Offset value for node.
WGXLS - Word-to-Grammar eXtra Layer Synapse. Weight based on probability of grammar node Y given that word node X is active.
GWXLS - Grammar-to-Word eXtra Layer Synapse. Weight based on probability of word node Y given that grammar node X is active.
~GWXLS - Not Grammar-to-Word eXtra Layer Synapse. Probability of word node Y given that grammar node X does is not active.
PAW - Phoneme Activation (PA)-to-Word. Weight based on the probability of word node Y given that the combined activation of word Y's phonemes is positive.
~PAW - Not PA-to-Word. Weight based on the probability of word node Y given that the combined activation of the word Y's phonemes is negative.
PA Bias - Bias for PA node
PPAXLS - Phon-to-PA Node eXtra Layer Synapse. Weight based on the presence of phoneme X in word Y
~PPAXLS - Not Phon-to-PA node eXtra Layer Synapse. Weight based on the absence of phoneme X in word Y
WPXLS - Word-to-Phon eXtra Layer Synapse. Weight base of probability of Phoneme Y given word X.