Appendix 1: Formal Language Definitions

Portions of this material are adapted from (Gurari 1989).

A Type 0 grammar ***G*** is defined as a quadruple <***A*** , Σ , ***P*** , ***S*** >, where

A is an alphabet, whose elements are called non-terminal symbols. ***A**** denotes the set of possible words formed from concatenating these symbols.

Σ is an alphabet disjoint from N, whose elements are called terminal symbols.

P is a relation of finite cardinality on (***A***)*, whose elements are called production rules. Moreover, each production rule (α,β) in P, denoted $\alpha\to\beta$, must have at least one non-terminal symbol in ***A*** . In each such production rule, α is said to be the left-hand side of the production rule, and β is said to be the right-hand side of the production rule.

S is a symbol in ***A*** called the start , or sentence , symbol.

A grammar ***G*** = < ***N, Σ, P, S*** > is said to be a right-linear grammar if each of its production rules is either of the form $\alpha\to x\beta$ or of the form $\alpha \to$ x, where α and β are non-terminal symbols in ***A*** and *x* is a string of terminal symbols in Σ*.

The grammar is said to be a left-linear grammar if each of its production rules is either of the form $\alpha\to\beta x$ or of the form $A\to x$, where α and β are nonterminal symbols in ***A*** and *x* is a string of terminal symbols in Σ*.

The grammar is said to be a regular grammar if it is either a right-linear grammar or a left-linear grammar. A language is a **regular language** if it is generated by a regular grammar.

Such a Type 0 grammar G = <***A, Σ, P, S***> is said to be context-free if each of its production rules has exactly one non-terminal symbol on its left hand side, that is, if each of its production rules is of the form A->alpha .

The concept of a grammar can be extended to a *stochastic* grammar. The set ***P*** of productions can be associated one to one with an set π of probabilities on each production. The resulting system <***A, Σ, P, π, S***> is a stochastic grammar. The set of productions is a stochastic language. A state transition graph consisting of nodes for each element in **A** can be defined, with labeled arcs giving the probabilities of emitting a new token; this graph specifies a Markov process which generates strings in the stochastic language ***L.***

Appendix 2: Discrete Markov Chains

In this appendix basic definitions related to Markov chains are presented following the treatment in the Sochastic Processes text of Lawler (Lawler 1995). I occasionally interject comments drawing connections to dynamical systems. While Markov chains share certain concepts with dynamics systems such as periodicities, transient states, the basic dichotomy between deterministic and probabilistic systems remains. By considering coarse grained partitions of the dynamical phase space(as is also performed for symbolic dynamics) as the states in a Markov chain, one can convert from the deterministic to the stochastic mode of analysis.

STOCHASTIC AND MARKOV PROCESSES

A **stochastic process** is a collection of random variables X_t indexed by time. When time is a subset of the nonnegative integers {0,1,2 ...} the process is called discrete time. The random variables take values in a state space; this may be discrete (a finite or countably infinite set) or continuous. A **Markov process** is a stochastic process with the restriction that the change at time t is determined by the value of the process (i.e. the value of the state space scalar or vector in R^d) at time t, and not by values at times before t.

MARKOV CHAINS

A time homogeneous Markov chain is a Markov process described by a initial probability distribution and a transition probability matrix P, where the elements in the matrix P_{ij} are independent of time. The matrix P must be a stochastic matrix, satisfying the conditions:

$$0 \le P_{ij} \le 1, 1 \le i, j \le N$$

$$\sum_{j=1}^{N} P_{ij} = 1, 1 \le i \le N$$

The n-step transition probabilities $p_n(i,j)$ are given by P^n.

An absorbing state is a state which leads to itself with probability one. This is equivalent to a fixed point attractor in a dynamical system.

A Markov chain is irreducible if all states communicate, i.e. there is a path between the two states in the transition matrix. Otherwise, the state space is partitioned into disjoint sets called communication classes. These may be transient or recurrent, inheriting properties from their constituent states. A transient state will leave the state with probability 1 (when the system is captured by an absorbing state).

The partitioning of states into transients and absorbing states corresponds to the partitioning of contracting dynamical systems into attractors and (basin) transients. For expanding (chaotic) dynamical systems, there may be forbidden regions of phase space on the attractor, which nevertheless could be given as an initial condition or reached by a perturbation.

A recurrent Markov chain is one for which each state is visited infinitely many times. In contrast, a transient chain is one for which each state is visited a finite number of times.

Appendix 3: Signal Analysis

Autocorrelation

Autocorrelation emphasizes periodic components of a time series by comparing values separated by a regular time interval (lag). Each sample of the time series is multiplied by the value shifted in time by a fixed lag; the sum of these products is the autocorrelation function for a particular lag time. The autocorrelogram is generated by

1. Removing the mean from the signal
2. Normalizing by signal power
3. Plotting lag time (x axis) vs. the normalized correlation coefficient.

A similar process, cross-correlation, is used to compare two time series. Autocorrelation can, of course, be viewed as a special case of cross correlation.

$$R_{xy}[k] = 1/N\sum_{n=0}^{N-1} x[n]y[n+k]$$

where x[n] and y[n] are values of time series x and y at time n, k is the lag of y[n] with respect to x[n], and N is the number of samples in series or window in which the function is computed.

Let N be the number of sampled points in two signals x and y.

Power Spectrum

$$S_{xx}(f) = \frac{FFT(x) * FFT^{*}(x)}{N}$$

where FFT^{*} is the complex conjugate of *FFT(x)* (e.g. the negated imaginary part of *FFT(x))*.

Cross Power Spectrum

$$S_{xy}(f) = \frac{FFT(y) * FFT^{*}(x)}{N^{2}}$$

Coherence function

$$C_{xy} = \frac{[|(Averaged\ S_{\mathrm{xy}}(f))|]^{2}}{Averaged\ \mathrm{S}_{\mathrm{xx}}(f) * Averaged\ \mathrm{S}_{\mathrm{yy}}(f)}$$

The coherence function requires an average of two or more measurements of the signals under analysis. For a single measurement, it would register unity at all frequencies. To average a complex quantity such as the cross power spectrum $S_{xy}\ (f)$, sum it in the complex form, divide by the number of averaging trials, then convert to magnitude and phase with rectangular to polar conversion.. The auto power spectra, $S_{AA}(f)$ and $S_{BB}(f)$ are real quantities.