Shift Space of First Order Language Symbols

Thobias Høivik

November 21, 2025

1 Cool Beans

Definition 1.1: First Order Language

A first-order language $\mathfrak{L}$ is a set of symbols composed of two disjoint subsets:

1. The first part, which is common to all languages, consists of "(" and ")" together with the following symbols: the variables $\mathcal{V} = \{v_n : n \in \mathbb{N}\}$, the equality symbol "=", connectives "¬", "∧" and the existential quantifier "∃".
2. The non-logical symbols, consisting of
   - a set of constant symbols $\mathscr{C}^{\mathscr{L}}$.
   - a sequence of sets $\mathscr{F}_n^{\mathfrak{L}}$, $n \in \mathbb{Z}^+$, where elements of this set are called n-ary functions symbols.
   - a sequence of sets $\mathscr{R}_n^{\mathfrak{L}}$, $n \in \mathbb{Z}^+$, where elements of this set are called n-ary relation symbols or predicates depending on the context.

Definition 1.2: The Language Alphabet

Let $\mathfrak{L}$ be a first-order language as defined above. We define the alphabet $\mathscr{A}_{\mathfrak{L}}$ of the dynamical system to be the complete set of symbols comprising the language $\mathfrak{L}$:

$$\mathscr{A}_{\mathfrak{L}} = \text{Set of all symbols in } \mathfrak{L}$$

Since $\mathfrak{L}$ contains a countably infinite set of variables $\mathcal{V}$ (and potentially infinite sets of non-logical symbols), the alphabet $\mathscr{A}_{\mathfrak{L}}$ is countably infinite. We equip $\mathscr{A}_{\mathfrak{L}}$ with the discrete topology τ_{disc}.

Definition 1.3: The Full Shift Space over $\mathfrak{L}$

The full shift space over the first-order language $\mathfrak{L}$, denoted $\Sigma_{\mathfrak{L}}$, is the set of all two-sided infinite sequences (or words) formed by symbols from $\mathscr{A}_{\mathfrak{L}}$:

$$\Sigma_{\mathfrak{L}} = \mathscr{A}_{\mathfrak{L}}^{\mathbb{Z}} = \{s = (\dots, s_{-1}, s_0, s_1, \dots) : s_i \in \mathscr{A}_{\mathfrak{L}} \text{ for all } i \in \mathbb{Z}\}$$

We equip $\Sigma_{\mathfrak{L}}$ with the product topology, where a basis for the open sets is given by the cylinder sets $C_W = \{s \in \Sigma_{\mathfrak{L}} : s_{[i,i+k-1]} = W\}$ for any finite word (block) W of length k occurring at position i.

Definition 1.4: The Shift Map

The shift map (or shift automorphism) $\sigma : \Sigma_{\mathfrak{L}} \to \Sigma_{\mathfrak{L}}$ is the function defined by shifting every sequence one position to the left:

$$\sigma(s)_i = s_{i+1} \quad \text{for all } s \in \Sigma_{\mathfrak{L}} \text{ and } i \in \mathbb{Z}$$

The pair $(\Sigma_{\mathfrak{L}}, \sigma)$ forms a topological dynamical system. Since the alphabet $\mathscr{A}_{\mathfrak{L}}$ is infinite, the space $\Sigma_{\mathfrak{L}}$ is not compact, which contrasts with classical symbolic dynamics over finite alphabets.

Definition 1.5: The Subshift of Well-Formed Terms

A subshift $X \subseteq \Sigma_{\mathfrak{L}}$ is a closed and σ-invariant subset of $\Sigma_{\mathfrak{L}}$. We can define a specific subshift X_{Term} by imposing a logical constraint:

$$X_{\text{Term}} = \{s \in \Sigma_{\mathfrak{L}} : \text{no block in } s$$
$$\text{is a finite sequence of symbols that violates the } \mathfrak{L}\text{-grammar for a well-formed term}\}$$

This definition means that every finite word W appearing in a sequence $s \in X_{\text{Term}}$ must be a $\mathfrak{L}$-grammatically valid string (not necessarily a complete term, but one that doesn't contain a forbidden block $\mathscr{F}$ that violates the rules of $\mathfrak{L}$-term formation, such as $x(\wedge)$. The set of forbidden blocks $\mathscr{F}$ is typically context-free or higher in complexity due to the recursive nature of term formation.