Theoretical Computer Science 567 (2015) 21–45

Contents lists available at ScienceDirect

Theoretical Computer Science

www.elsevier.com/locate/tcs

Category theory of symbolic dynamics ☆

Ville Salo, Ilkka Törmä *

TUCS – Turku Centre for Computer Science, Department of Mathematics and Statistics, 20014 University of Turku, Finland

ARTICLE INFO

Article history:
Received 11 September 2013
Received in revised form 12 June 2014
Accepted 15 October 2014
Available online 22 October 2014
Communicated by G. Rozenberg

Keywords:
Symbolic dynamics
Category theory
Subshift

ABSTRACT

We study the central objects of symbolic dynamics, that is, subshifts and block maps, from the perspective of basic category theory, and present several natural categories with subshifts as objects and block maps as morphisms. Our main goals are to find universal objects in these symbolic categories, to classify their block maps based on their category theoretic properties, to prove category theoretic characterizations for notions arising from symbolic dynamics, and to establish as many natural properties (finite completeness, regularity etc.) as possible. Existing definitions in category theory suggest interesting new problems in symbolic dynamics. Our main technical contributions are the solution to the dual problem of the Extension Lemma and results on certain types of conserved quantities, suggested by the concept of a coequalizer.

1. Introduction

Like many branches of mathematics, symbolic dynamics is the study of a category: the category with subshifts as objects and block maps as morphisms. Of particular interest is the case where the objects are SFTs or sofic shifts, and we will mostly concentrate on such subcategories. Like in many branches of mathematics, the first question that comes to mind is still open: 'When are two objects isomorphic?', and a lot of mathematics has been developed trying to answer this question [1]. Questions such as 'When is an object a subobject of another?' (the embedding problem) and 'When does an object map onto another?' (the factoring problem) have been solved at least in important special cases (see Factor Theorem and Embedding Theorem in [2], and the article [3]).

A nice feature of category theory is that it allows one to define notions such as 'isomorphism', 'embedding' and 'factoring' without referring to anything except morphisms between objects. For example, two objects X and Y are isomorphic, in the sense of category theory, if there exist morphisms $f : X \to Y$ and $g : Y \to X$ such that $g \circ f = \mathrm{id}_X$ and $f \circ g = \mathrm{id}_Y$. Usually, this turns out to be the 'correct' notion for isomorphism. For embedding (and factoring), the situation is more complicated: often the most natural definition of an embedding is that it is injective. The notion of injectivity (and surjectivity) is, however, impossible to define categorically, since a category does not know what its morphisms actually are: they need not even be functions! Because of this, multiple categorical variants of injectivity have been defined, including monicness, split monicness and regular monicness. These are generalizations of the various ways in which an injective map should behave in relation to other morphisms. In sufficiently nice categories (for example, the category of sets), these all correspond to injectivity, but in many categories, they state some other property, and often raise new natural questions.

☆ Research supported by the Academy of Finland Grant 131558.

* Corresponding author. Tel.: +358 2 333 5056.
E-mail addresses: vosalo@utu.fi (V. Salo), iatorm@utu.fi (I. Törmä).

http://dx.doi.org/10.1016/j.tcs.2014.10.023
0304-3975/

Because categorical notions depend on *all* morphisms of the category, we define several categories (thirteen, to be exact) whose objects are subshifts and morphisms are block maps between them. We begin our study in Section 4 by considering the various categorical definitions of injectivity and surjectivity for our categories. We see that these categories are not nearly as 'nice' as that of sets, in that usually these notions do not characterize injectivity or surjectivity in our categories. Interestingly, the characterization of split monicness comes from the well-known Extension Lemma. The dual concept of split epicness turns out interesting as well, and we find a characterization in the SFT case. Our discussion of morphisms is motivated as the study of how well standard notions of category theory describe the world of block maps, and in Section 5, we ask a kind of converse question of whether category theory can describe standard notions of symbolic dynamics. Here, we mostly concentrate on the properties of objects.

In Section 6, we move on to important category theoretical closure properties: the existence of (finite) limits and colimits of diagrams. The importance of these notions is that many category theoretical definitions are just limits of diagrams of certain types. The case of coequalizers turns out to be the most intricate, and we present an undecidability result related to it. In Section 7, we discuss the existence of images, disjoint unions and quotients in a categorical sense, that is, whether our categories are regular, coherent and exact, respectively.

The topics of cellular automata (particular kinds of block maps) and category theory have been previously explicitly discussed together at least in [4], but the approach there is very different. There, cellular automata are constructed using category theoretical tools, and properties that 'come for free' from category theoretical generalities are investigated.

2. Definitions and notation

In this section, we establish the basic terminology and notation used in this article. As we are working in the intersection of two quite distinct fields, symbolic dynamics and category theory, we try to be as complete as possible.

2.1. Symbolic dynamics

For a finite set S (an *alphabet*) with the discrete topology, we denote by $S^{\mathbb{Z}}$ the space of two-way infinite *configurations* (or *points*) over S with the product topology, and call it the *full shift on* S. The *shift action* $\sigma : S^{\mathbb{Z}} \to S^{\mathbb{Z}}$ is defined by $\sigma(x)_i = x_{i+1}$ for all $i \in \mathbb{Z}$. A fixed point of σ is called a *uniform point*. A closed subset X of a full shift with $\sigma(X) = X$ is called a *subshift*. We say that a word $w \in S^*$ occurs in $x \in S^{\mathbb{Z}}$ if $x_{[i,i+|w|-1]} = w$ for some $i \in \mathbb{Z}$, and also write $w \sqsubset x$. An alternative characterization of subshifts is by means of *forbidden words*: $X \subset S^Z$ is a subshift if and only if there exists a set of words $F \subset S^*$ such that $X = \{x \mid \forall w \in F : w \not\sqsubset x\}$. If the set of forbidden words can be taken to be finite, X is called a *subshift of finite type*, or SFT for short, and if the set is a regular language, X is called *sofic*. If X is an SFT and Y is any subshift, then $X \cap Y$ is a *subSFT* of Y. If the SFT (or subSFT of Y) X can be defined by forbidden words of length at most m (in addition to the forbidden words of Y), we say m is a *window size* of X (relative to Y). We denote by σ_X the restriction of σ to X.

The words of length n appearing in configurations of X are denoted $\mathcal{B}_n(X)$, and we denote the *language* of X by $\mathcal{B}(X) = \bigcup_{n\in\mathbb{N}} \mathcal{B}_n(X)$. Since a subshift is defined by its language [2], we may also denote $X = \mathcal{B}^{-1}(L)$, if $L \subset S^*$ is an extendable language (for every $v \in L$ there exist $u, w \in S^+$ with $uvw \in L$) such that $\mathcal{B}(X)$ is the language of subwords of words in L. Usually, when using this notation, we write a regular expression in place of L. An SFT $X \subset S^{\mathbb{Z}}$ can also be defined by giving a set of *allowed words* $A \subset S^n$ for some $n \in \mathbb{N}$ such that $X = \{x \in S^{\mathbb{Z}} \mid \forall i : x_{[i,i+n-1]} \in A\}$.

Example 1. The sofic shift $\mathcal{B}^{-1}(0^*10^*)$ consists of exactly those configurations of $\{0,1\}^{\mathbb{Z}}$ that contain at most one 1. The subshift $\mathcal{B}^{-1}((0^*10)^*)$ is an SFT, and can be defined by the single forbidden word 11. For a fixed $p \in \mathbb{N}$, the SFT $\mathcal{B}^{-1}((0^{p-1}1)^*)$ contains exactly p points with spatial period p, and we use it as a 'canonical' p-periodic subshift. As a dynamical system, it is isomorphic to $(\mathbb{Z}_p, n \mapsto n+1 \bmod p)$, the set of integers modulo p with incrementation.

For two subshifts $X \subset S^{\mathbb{Z}}$ and $Y \subset R^{\mathbb{Z}}$, define $X \times Y \subset (S \times R)^{\mathbb{Z}}$ as the coordinatewise product

$$\left\{z \in (S \times R)^{\mathbb{Z}} \mid \ldots p_1(z_{-1})p_1(z_0)p_1(z_1)\ldots \in X \wedge \ldots p_2(z_{-1})p_2(z_0)p_2(z_1)\ldots \in Y\right\},$$

where p_1 and p_2 are the appropriate projections from $S \times R$. Define also $X \dot\cup Y$ as their symbol-disjoint union, where we replace S and R with disjoint sets if necessary.

The *syntactic monoid* $\mathrm{Syn}(X)$ of a subshift $X \subset S^{\mathbb{Z}}$ is defined as $S^*/{\sim_X}$, where $u \sim_X v$ denotes that $wuw' \in \mathcal{B}(X)$ if and only if $wvw' \in \mathcal{B}(X)$ for all $w, w' \in S^*$. We also denote $(v)_X = v/{\sim_X}$. It is known that sofic shifts are exactly those subshifts whose syntactic monoid is finite.

A *block map* is a continuous function $f : X \to Y$ from a subshift to another with $f \circ \sigma_X = \sigma_Y \circ f$. Block maps are defined by *local functions* $F : \mathcal{B}_{2r+1}(X) \to \mathcal{B}_1(Y)$ by $f(x)_i = F(x_{[i-r,i+r]})$, where $r \geq 0$ is called a *radius* of f. If f is surjective, Y is a *factor* of X, and if it is bijective, X and Y are *conjugate*. The block map itself is called an *embedding*, a *factor map* or a *conjugacy*, if it is injective, surjective or bijective, respectively. We say f is *preinjective* if $f(x) \neq f(y)$ whenever $x \neq y \in X$ are *asymptotic*, that is, they differ in finitely many coordinates. If $X = Y$, then f is called a *cellular automaton* on X.