Theoretical Computer Science ••• (••••) •••–•••

Contents lists available at ScienceDirect

Theoretical Computer Science

www.elsevier.com/locate/tcs

Dipole codes attractively encode glue functions

Dhananjay Ipparthi [a], Massimo Mastrangeli [b], Andrew Winslow [c,*]

[a] IRIDIA, CoDE, Université libre de Bruxelles, Brussels, Belgium
[b] Max Planck Institute for Intelligent Systems, Stuttgart, Germany
[c] Département d'Informatique, Université libre de Bruxelles, Brussels, Belgium

ARTICLE INFO

Article history:
Received 11 August 2015
Received in revised form 28 November 2015
Accepted 11 February 2016
Available online xxxx

Keywords:
Self-assembly
Mesoscale
Magnetic

ABSTRACT

Dipole words are sequences of magnetic dipoles, in which alike elements repel and opposite elements attract. Magnetic dipoles contrast with more general sets of bonding types, called *glues*, in which pairwise bonding strength is specified by a *glue function*. We prove that every glue function g has a set of dipole words, called a *dipole code*, that *attractively encodes* g: the pairwise attractions (positive or non-positive bond strength) between the words are identical to those of g. Moreover, we give such word sets of asymptotically optimal length. Similar results are obtained for a commonly used subclass of glue functions.

1. Introduction

Self-assembly is the autonomous organization of components into structures without supervision [1]. Here we consider controlling self-assembly using fixed arrangements of magnetic dipoles, specifically leveraging their ability to attract and repel according to their spatial configurations. Some previous microscale [2–4] and mesoscale [5–7] self-assembling systems have used the capillary effects of surface tension as alternative bonding mechanisms. More recently, molecular recognition of mesoscale components via surface chemistries has also been used [8–10]. However, magnets are among the most common sources of interaction force in micro- and mesoscale assembly, and they are used in both *active* components that change bonding behavior [11–13] and in *passive* components whose behavior is fixed [14–17].

A primary limitation of dipole-based bonding is the limited number of interactions between dipoles: alike poles repel, while opposite poles attract. Frameworks by Bhalla et al. [18,14] and Majumder and Reif [19] describe an approach for obtaining more complex behaviors by arranging sequences of dipoles along the boundaries of components. These *dipole codes* are used to obtain many distinct bonding sites, called *glues*, that interact only with a unique complementary glue. Bhalla et al. [18,14,20] have demonstrated that dipole codes also work experimentally.

The use of *DNA codes* is ubiquitous in DNA-based nanoscale self-assembly, where sequences of repeating nucleotides from the alphabet $\{\mathtt{A}, \mathtt{T}, \mathtt{C}, \mathtt{G}\}$ have been used experimentally to form dozens [21–23] or even hundreds [24,25] of glues. Their theoretical study also is extensive, as seen in several surveys [26–30]. However, there are several practical aspects that differentiate the design of DNA codes from dipole codes.

For instance, the elasticity of DNA requires that codes must disallow multiple portions of a single code to bond [31,32], while chemistry requires that the codes must have balanced occurrences of letters $\mathtt{A}$, $\mathtt{T}$ and $\mathtt{G}$, $\mathtt{C}$ [33–35]. On the other hand, dipole codes have only a single pair of bonding letters, and inconsistency of mesoscale mixing allows codes to bond with even a single dipole pair [14].

* Corresponding author.
E-mail addresses: dhananjay.ipparthi@ulb.ac.be (D. Ipparthi), mastrangeli@is.mpg.de (M. Mastrangeli), andrew.winslow@ulb.ac.be (A. Winslow).

http://dx.doi.org/10.1016/j.tcs.2016.02.038
0304-3975/

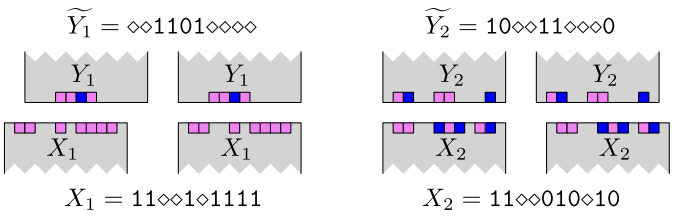

Fig. 1. Left: the pair $X_1, \widetilde{Y_1}$ is attracted, since $f(X_1, \widetilde{Y_1}) = 1 > 0$, and aligned. Right: the pair $X_2, \widetilde{Y_2}$ is not attracted, since $f(X_2, \widetilde{Y_2}) = -3 + 2 = -1 \not> 0$, and not aligned, since $f(X_2[..-2], \widetilde{Y_2}[2..]) = 3 > 0$.

Theoretical work on DNA-based systems has also demonstrated that the addition of a repelling force to systems with many glues increases computational power [36,37] and efficiency [38,39]. Thus the construction of large numbers of glues with magnetic dipole sequences gives access to yet more techniques for controlling assembly.

Our contribution The frameworks of both Bhalla et al. [18,14] and Majumder and Reif [19] have several drawbacks. First, neither formalizes how to obtain sets of magnetic dipole sequences that encode the behavior of a desired number of glues, nor glues with complex pairwise interactions. Second, their sequences require precise control of the system's *temperature*, the amount of force necessary for a bond to be irreversible, as well as additional component geometry. Both are needed to prevent undesired bonds caused by dipole sequence pairs that only partially match or are misaligned.[1] Our work addresses both drawbacks, giving magnetic dipole sequences that encode pairwise bonding behaviors of arbitrarily many glues at fixed temperature and without misalignment.

Section 2 begins by giving a formal model of magnetic dipole sequences, called *dipole words*, and the net forces between them. The pairwise bonding behaviors of a set of glues are defined by a *glue function*, and we define what it means for a set of dipole words to *encode* a glue function and thus be a *dipole code*. These definitions allow for the possibility of misaligned or weak bonds. They require that the dipole words work at fixed temperature (a pair can bond if the net attractive force is positive) and that all misaligned bonds are non-attractive.

Section 3 contains a "warmup" dipole word set construction that encodes *canonical* glue functions, where each glue only bonds to itself or to a unique complementary glue. For any such function over k glues, this construction gives a dipole code of length $O(k)$ that encodes it.

Section 4 improves this construction by extending it to all glue functions, allowing for *flexible glues* (see [40]) that bond to many others. Section 5 improves on length of dipole codes for canonical glue functions only, obtaining length-$O(\log k)$ codes. For both of these results, we also prove that the word lengths of the second and third constructions are asymptotically optimal. Finally, Section 6 poses several remaining open problems.

2. Definitions

Letters and words A *letter* is a symbol x in the *alphabet* $\Sigma = \{0, 1, \diamond\}$. A *dipole word* or simply a *word* is a sequence of letters, and the *length* of a word W, denoted $|W|$, is the number of letters in W. For an integer $i \in \{1, 2, \ldots, n\}$, $W[i]$ refers to the ith letter of W and $W[-i]$ refers to the ith from the last letter of W. The *reverse* of a word W, written $\widetilde{W}$, is the letters of W in reverse order.

A *subword* of W is a contiguous sequence of letters in W. For integers $1 \le i, j \le |W|$ with $i \le j$, $W[i..j]$ denotes the subword of W from $W[i]$ to $W[j]$, inclusive. As shorthand, $W[-i..] = W[-i..|W|]$ and $W[..j] = W[1..j]$.

Forces The *force* of a pair of letters x, y is defined by function

$$f(x, y) = \begin{cases} 1 & : \{x, y\} = \{0, 1\} \\ -1 & : \{x, y\} \in \{\{0\}, \{1\}\} \\ 0 & : \text{otherwise} \end{cases}$$

and a pair of letters x, y is *attracted* or *repelled* provided $f(x, y) = 1$ or $f(x, y) = -1$, respectively. Similarly, for words X, Y with $|X| = |Y|$, the *force* of the pair X, Y, denoted $f(X, Y)$, is defined as

$$f(X, Y) = \Sigma_{i=1}^{|X|} f(X[i], Y[i])$$

and the pair X, Y is *attracted* provided $f(X, Y) > 0$ (see Figure 1).

[1] Section 5.2 and Figure 11 of [18] and Section 4 of [19] discuss these difficulties.

Please cite this article in press as: D. Ipparthi et al., Dipole codes attractively encode glue functions, Theoret. Comput. Sci. (2016), http://dx.doi.org/10.1016/j.tcs.2016.02.038