Information Sciences 208 (2012) 39–54

Contents lists available at SciVerse ScienceDirect

Information Sciences

journal homepage: www.elsevier.com/locate/ins

# Local measures of information storage in complex distributed computation

Joseph T. Lizier [a,b,c,*], Mikhail Prokopenko [b], Albert Y. Zomaya [c]

[a] Max Planck Institute for Mathematics in the Sciences, Inselstraße 22, 04103 Leipzig, Germany
[b] CSIRO Information and Communications Technology Centre, PO Box 76, Epping, NSW 1710, Australia
[c] School of Information Technologies, The University of Sydney, NSW 2006, Australia

ARTICLE INFO

*Article history:*
Received 2 June 2011
Received in revised form 12 March 2012
Accepted 13 April 2012
Available online 27 April 2012

*Keywords:*
Information storage
Intrinsic computation
Complex systems
Information theory
Cellular automata
Particles

ABSTRACT

Information storage is a key component of intrinsic distributed computation. Despite the existence of appropriate measures for it (e.g. excess entropy), its role in interacting with information transfer and modification to give rise to distributed computation is not yet well-established. We explore how to quantify information storage on a local scale in space and time, so as to understand its role in the dynamics of distributed computation. To assist these explorations, we introduce the active information storage, which quantifies the information storage component that is directly in use in the computation of the next state of a process. We present the first profiles of local excess entropy and local active information storage in cellular automata, providing evidence that blinkers and background domains are dominant information storage processes in these systems. This application also demonstrates the manner in which these two measures of information storage are distinct but complementary. It also reveals other information storage phenomena, including the misinformative nature of local storage when information transfer dominates the computation, and demonstrates that the local entropy rate is a useful spatiotemporal filter for information transfer structure.



## 1. Introduction

Information storage is considered an important aspect of the dynamics of many natural and man-made processes, for example: in human brain networks [22,46] and artificial neural networks [4], synchronisation between coupled systems [5,42], coordinated motion in autonomous [1] and modular robots [43], attractor dynamics of cellular automata [37], and in the dynamics of inter-event distribution times [13]. Despite the existence of suitable quantitative measures for information storage (e.g. the excess entropy [8]), the term is still often used rather loosely or in a qualitative sense. A major factor here is that there is not yet a well-established quantitative understanding of how information storage dynamically interacts with information transfer and modification to give rise to intrinsic distributed computation in multivariate systems.

In this paper we explore methods to quantify information storage in distributed computation. In particular, we focus on how information storage can be quantified on a *local scale* in space–time, which allows us to directly investigate the role of information storage in the *dynamics* of distributed computation.

We focus on cellular automata (CAs) which, as described in Section 2, are a popular model of distributed computation in which the notion of information storage is qualitatively well-understood. Indeed, the emergent structures in CAs known as "blinkers" have been conjectured to implement information storage, and we hypothesise that appropriate quantification of the dynamics of information storage will align with this conjecture.

* Corresponding author at: Max Planck Institute for Mathematics in the Sciences, Inselstraße 22, 04103 Leipzig, Germany.
*E-mail address:* lizier@mis.mpg.de (J.T. Lizier).

0020-0255/$ - see front matter © 2012 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.ins.2012.04.016

We define the concept of information storage as *the information in an agent, process or variable's past that can be used to predict its future*. We consider such storage in the *intrinsic computation* of the unfolding of a system's dynamics [11], which takes place regardless of whether it is explicitly interpretable as a computation to an observer. We describe in Section 4 how *total* storage *relevant to the future* of a process is captured by the existing measure *statistical complexity* [9], while the *total* storage *actually used* in the future of a process is captured by the existing measure *excess entropy* [8]. We then introduce active information storage in Section 5 to capture the amount of storage that is *currently in use* by a process. Our perspective of distributed computation is important, providing the perspective that information can not only be stored internally by an agent or element implementing the computation, but also stored in its environment for later retrieval.

We present the first application of local profiles of the excess entropy and the active information storage to cellular automata in Section 6. As hypothesised above, these applications provide the first quantitative evidence that blinkers are the dominant information storage entities there. This result is significant in marrying these quantitative measures of information storage with the popularly-understood qualitative notion of its embodiment in distributed computation. These measures also reveal other important local information storage phenomena, including the misinformative nature of storage when information transfer dominates the computation. Our application also demonstrates the manner in which these two measures of information storage are distinct but complementary, revealing different aspects of the information dynamics. Finally, we demonstrate that the entropy rate, a complementary measure to the active information storage, is a useful filter for viewing moving particle structures in CAs.

## 2. Cellular automata

Information storage has been a topic of interest in the context of distributed computation in *cellular automata* (CAs). As we will describe, the notion of information storage is qualitatively well-understood with regard to CAs, and as such we choose these as our application domain.

CAs are discrete dynamical lattice systems involving an array of cells which synchronously update their states as a homogeneous deterministic function (or rule) of the states of their local neighbours [51]. Here we will use Elementary CAs (ECAs), which consist of a one-dimensional array of cells with binary states, with each updated as a function of the previous states of themselves and one neighbour either side (i.e. neighbourhood size 3 or range $r = 1$). CAs are widely used to study complex computation, since certain rules (e.g. ECA rules 110 and 54, defined using the Wolfram numbering scheme [51]) exhibit *emergent coherent structures* which are not discernible from their microscopic update functions but which provide the basis for understanding the macroscopic computations being carried out [39,52].

These emergent structures are known as *particles*, *gliders* and *domains*. A domain may be understood as a set of background configurations in a CA, any of which will update to another such configuration in the absence of a disturbance. The most simple domain types involve periodic repetition of cell states in time and space. Domains are formally defined within the framework of computational mechanics [18] as spatial process languages in the CA. *Particles* are qualitatively considered to be moving elements of coherent spatiotemporal structure, in contrast to or against a background domain. *Gliders* are particles which repeat periodically in time while moving spatially, while stationary gliders are known as *blinkers*. Formally, particles are defined as a boundary between two domains [18]; as such, they can also be termed as *domain walls*, though this is typically used with reference to aperiodic particles (e.g. those in rule 18).

There are several long-held conjectures regarding the role of these emergent structures in the intrinsic distributed computation in CAs; i.e. how the cells process information in order to determine their collective future state [25,39]. Blinkers are generally held to be the dominant information storage elements, since local pattern maintenance is an information storage process. In contrast, particles are held to be the dominant information transfer entities, since they communicate coherent information about the dynamics in one area of the system to another (indeed, we have provided the first direct quantitative evidence for this conjecture with a measure of local information transfer in [31]). Studies of the density classification task with rule $\phi_{par}$ [39–41] help our intuition here. They suggest a human-understandable computation, with stationary blinkers used to store information about the local density of "1"'s in one region of the CA, while moving gliders are used to transfer information about these local densities between regions.

The presence of such emergent structure in CAs is revealed by *filtering* techniques, which highlight particles against the background domain. Early methods were hand-crafted for specific CAs (relying on the user knowing the pattern of background domains) [14,17], while later methods can be automatically applied to any given CA. These include: finite state transducers to recognise the regular spatial language of the CA using $\epsilon$-machines [18,19]; local information (i.e. local spatial entropy rate) [20]; the display of executing rules with the most frequently occurring rules filtered out [52]; and local statistical complexity (via a light-cone formulation) and local sensitivity [45].

Certainly, filtering is not a new concept, however the ability to *separately* filter different computational features is novel. For example, we have previously used local information *transfer* to filter *moving* particle structures in CAs [31], and a measure of local information *modification* to filter particle *collisions* in [32]. Here, we *hypothesise* that local measures of information *storage* should be useful filters for blinker structures: this would provide the first quantitative evidence for their computational role as dominant information storage structures in CAs. Together with our work regarding local information transfer [31] and information modification [32], the investigation of local information storage will show how these operations interrelate to give rise to complex behaviour, in comparison to other filters which give only a single view of where that