(anti-Ωx × Σz)-Based k-Set Agreement Algorithms

Zohir Bouzid1,* and Corentin Travers2,**

1 University Pierre et Marie Curie - Paris 6, LIP6-CNRS 7606, France
   zohir.bouzid@lip6.fr
2 LaBRI University Bordeaux 1
   travers@labri.fr

Abstract. This paper considers the k-set agreement problem in a crash-prone asynchronous message passing system enriched with failure detectors. Two classes of failure detectors have been previously identified as necessary to solve asynchronous k-set agreement: the class anti-leader anti-Ωk and the weak-quorum class Σk. The paper investigates the families of failure detector (anti-Ωx)1≤x≤n and (Σz)1≤z≤n. It characterizes in an n processes system equipped with failure detectors anti-Ωx and Σz for which values of k, x and z k-set-agreement can be solved. While doing so, the paper (1) disproves previous conjectures about the weakest failure detector to solve k-set-agreement in the asynchronous message passing model and, (2) introduces the first indulgent algorithm that tolerates a majority of processes failures.

Keywords: Set-agreement, asynchrony, failure detectors, indulgent algorithms.

1 Introduction

The k-set-agreement problem k-set-agreement [10] is one of the fundamental problem in fault tolerant distributed computing. In this problem, n processes starting each with an initial private value are required to agree on at most k values chosen among their initial values. The problem generalizes the consensus problem, which corresponds to the case where k = 1. In an asynchronous system, it is well known that 1-set-agreement is impossible as soon as at least one process may fail by crashing [17], whereas the case k = n does not require any coordination at all. For intermediate values of k (1 < k < n), asynchronous k-set agreement tolerating t crash failures is possible if and only if k > t [6][25][30].

Failure detectors. A failure detector is a distributed oracle that provides processes with possibly unreliable information on failures [9]. According to the quality

* Supported by DIGITEO project PACTOLE and by the ANR projects SHAMAN and ALADDIN.
** On leave from University Pierre et Marie Curie - Paris 6. Supported by the ANR project SPREADS and by the INRIA project REGAL.
of the information, several classes of failure detectors can be defined. Starting with [27,31], the failure detector approach has been investigated to alleviate the \(k\)-set-agreement impossibility in asynchronous systems. An algorithm that tolerates unreliable failure detection is said to be indulgent towards its failure detector [10,21]. Informally, an indulgent algorithm is always safe: it never violates the safety part of the problem it is supposed to solve, even when the underlying failure detector gives false information about failures.

The quest for the weakest failure detector for \(k\)-set-agreement. Given a distributed problem \(P\), a natural question is to determine the weakest failure detector for \(P\), that is a failure detector \(D\) which is both sufficient to solve the problem – there is an asynchronous algorithm based on \(D\) that solves \(P\) – and necessary, in the sense that any failure detector \(D'\) that allows solving \(P\) can be used to emulate \(D\).

The question of the weakest failure detector class for \(k\)-set agreement \((1 < k < n)\) has been first stated in [29]. This line of research [11,12,20,24] culminated with the work of Zieliński who established that the failure detector class \(\text{anti}\-\Omega^{n-1}\) is the weakest to solve \((n-1)\)-set-agreement in the wait-free shared memory model [32]. This has later been generalized to any \(k, 1 \leq k < n\) by three independent groups [2,15,18]. Informally, a failure detector \(\text{anti}\-\Omega^k\) outputs sets of \(n - k\) process ids such that some non faulty process id eventually never appear in the outputs.

The situation is different in the message passing model where the answer is known only for the two boundaries cases, i.e., \(k = 1\) (consensus) and \(k = n - 1\) [14]. For consensus \((k = 1)\), it has been shown that the class of eventual leader failure detector \(\Omega = \text{anti}\-\Omega^1\) is the weakest failure detector in the asynchronous message passing model in which a majority of processes are non-faulty \((t < \frac{n}{3})\) [8]. This result is generalized to the wait-free environment in [13] where it is shown that \(\Omega \times \Sigma\) is the weakest failure detector class for consensus when \(t < n\). Intuitively, failure detector \(\Sigma\) provides a reliable quorum system: when queried, a failure detector of the class \(\Sigma\) returns a sets of processes ids, such that (1) any two sets intersect and (2) eventually, every set contains only ids of correct processes. Actually, \(\Sigma\) is the weakest failure detector to implement a register in the message passing model [5,13].

Recently, the failure detector family \((\Sigma_k \times \Omega^k)_{1 \leq k < n}\) has been conjunctured to be the weakest failure detector classes for \(k\)-set-agreement [4]. Failure detector \(\Sigma_k\) and \(\Omega^k\) generalizes the classes \(\Sigma\) and \(\Omega\) respectively. Intuitively, a failure detector \(\Sigma_k\) allows up to \(k\) partitions: any collection of \(k + 1\) sets outputs by the failure detector contain at least two intersecting sets. \(\Omega^k\), which has been introduced by Neiger [28], outputs sets of \(k\) ids that eventually converge to a set including the id of a non-faulty process. It is shown in [4] that \(\Sigma_{n-1} \times \Omega^{n-1}\) is equivalent to the loneliness failure detector \(\mathcal{L}\) which is the weakest failure detector class for \((n-1)\)-set-agreement [14]. Before this paper, nothing specific was known about the power of \(\Sigma_x \times \Omega^x\) to solve \(k\)-set-agreement, for \(1 < x < n-1\).