Proceedings of the 26th Conference
on Decision and Control
Los Angeles, CA • December 1987

Distributed Asynchonous Detection: General Models

John S. Baras*
Electrical Engineering and Systems Research Center
University of Maryland, College Park MD 20742

Abstract

We introduce axiomatically, new non-commutative probability models for the data in distributed detection asynchronous schemes. We completely characterize mathematically these models and their representation. They represent extensions of models for communicating asynchronous processes. We provide a general theorem characterizing the representation of data collected by a distributed asynchronous detection scheme. We discuss how this model can be used to obtain performance bounds.

Summary

In [1] we proposed new non-commutative probability models for multi-agent stochastic control problems. Our interest stems primarily from long range objectives to develop theories and algorithms that can properly incorporate the following: asynchrony between agents, anticipatory phenomena, interaction between information and control, local state models supported by locally collectable data, sensor "domain", controller "region of effectiveness", duality of information and control, sensor "fusion" or coordination.

In the present paper we provide summary of our results on the problem of Distributed M-array Detection with asynchronous operation. Basically there are M hypothesis $H_1, H_2, \dots, H_M$ affecting the system and N agents $A_1, \dots, A_N$. Agents collect data, communicate and make inferences. Regarding communication we develop a framework that permits information constraints, such as capacities, etc, and not constraints in terms of what each agent computes and communicates. The "fusion" strategy or "communication" strategy emerges from the mathematical framework. Due to space limitations we provide a summary only here. Detailed discussion and proofs can be found in [2].

Beginning from fundamental requirements on the data and propositions that appear in distributed detection we first develop some algebraic structures. First, a *simple proposition* or *simple event* is a proposition that can admit a yes (usually assigned the binary value 1) or no (assigned the value 0) answer only, regarding their validity. Their validity can be verified (ascertained) by some combination of the data (measurements, experiments) performed by the various agents. We denote by E the set of simple events (or propositions).

It is important to note that "ambiguous" events (i.e., requiring probability assignments for their validity) are not simple events. They are constructed later in our theory. There is a set of natural axioms we impose on E, supported by databases operating in multi-sensor, distributed stochastic systems. We also have two important operations of implication (denoted by $\leq$) and orthocmplementation (denoted by $'$). We then have:

Theorem 1: The set of simple propositions, in a distributed, asynchromous stochastic system is an *orthomodular σ-orthoposet.*

This however is not a complete characterization. The reason is that the data bases of such a system cannot just be characterized by the *logic* (or logical structure) of the simple events that can be verified by the agents. The structure of the "logic" by itself is not sufficient to determine the mathematical formalism which should be employed. It is a fact that a mathematical theory of distributed detection (or estimation) (and more generally, of stochastic large scale systems) is used not so much to reproduce the logical properties of simple yes-no experiments performed and answerable by the agents, but rather to compute statistics of "system state" transitions and of outcomes of more complicated experiments (measurements). Therefore, we next unify the probabilistic aspects with the logical aspects.

For the treatment of distributed detection and estimation problems, it suffices to introduce two more elements in the picture. Thus we are led to consider event-state structures. We think of states as the set of all possible (or just pertinent to the problem) configurations. We emphasize that we do not assume a memory interpretation for states.

We consider both global and local *event-state structures.*

Under certain natural anioms we show the following.

Theorem 2: Given an event-state structure (E, S, P) satisfying certain axioms, compatible with data bases in multi-agent systems, one can construct $(E, \leq, ')$ and S, *s.t.*

(a) $(E, \leq, ')$ is an orthomodular σ-poset
(b) $\hat{S}$ is a strongly order-determining σ-convex set of probability measures on $(E, \leq, ')$
(c) $a \mapsto \mu_a$ is a bijection of S onto $\hat{S}$.

There is actually a converse, asserting that the above representation is "faithful", in the sense that it can generate the statistics on which it was based.

In establishing this results a whole sequence of important constructs and intermediate results are obtained. We list some here: The states in $\hat{S}$ are basically *probability assignments* to simple events; construction of "mixture" states; prior probability about "states"; minimality of the state set supported by the observations; local states.

* The work of this author was supported partially through NSF Grant NSFD CDR-85-00108 and partially through ONR Grant N00014-83-K-0731.

There are two generic examples of such event-state structures. In the first we consider $P(\mathcal{H})$, the set of all gonal projections on a separable complex Hilbert space $\mathcal{H}$, and let $\leq$ be the usual order of projections. Let P' be the orthogonal complement of P. Then $(P(\mathcal{H}), \leq, ')$ is an orthomodular σ–orthoposet. Let S be the set of all positive, trace class, self-adjoint operators on $\mathcal{H}$, with trace one. Let $\hat{S} = \{\mu_\rho(\cdot), \rho \in S; \mu_\rho(P) = Tr[\rho P]\}$. Then $\hat{S}$, $P(\mathcal{H})$ form an event-state structure as discussed here. The second example, consists of a σ–algebra E of subsets of a set X, and a σ–convex, strongly order-determining set $\hat{S}$ of probability measures on X. This is the classical Kolmogorov model of probability theory with several probability measures.

The next important issue deals with communication constraints, incompatible events and event-state-operation structures. On intuitive grounds, we expect that in a multi-agent system incompatible events appear; that is, events such that their occurrence cannot be simultaneously verified by two or more agents. This can be seen as a manifestation of communication constraints in a distributed sensor network, for example.

Consider for a moment a distributed sensor network. It is clear that a specific sensor will be able to verify the occurrence or not of a restricted set of simple events. One may legitimately define this subset as the *domain of observation* or *sensor range* of the sensor. Similarly, an agent in a multi-agent stochastic control problem will be able to influence the occurrence of a subset of simple events. One may define this subset as *the domain of influence* or *control range* of the agent.

We need to develop new probabilistic models that incorporate such concepts in them. We do this by introducing a generalization of "conditioning", which we call *operation.* We construct an *event-state operation* structure: (E, S, P, T). Here (E, S, P) is an event-state structure and

$$T: \qquad E \longrightarrow \Sigma = \{\text{set of all maps from S to S}\}$$

We impose natural axioms on T. In particular we identify a natural association of simple events with a subset of operations. For a simple event p, let T_p represent the corresponding operation. If x is a state then $T_p\, x$ is the *new state* conditioned on occurrence of the event p and prior state x. Let

$$\Sigma_T = \{T_{p_1}\, o\, T_{p_2}\, o \cdots T_{p_n}; p_1 \cdots p_n \in E\}$$

be the set of *operations.* (Σ_T, o) is a multiplicative subsemigroup of (Σ, o). We introduce the * operation on Σ_T, to mean reversal of application; it is an involution. We can then show

Theorem 3: If (E, S, P, T) is an event- state-operation structure supported by the databases and conditioning of a distributed asynchronous stochastic system, then $(\Sigma_T, 0, *, \sim)$ is a Baer*-semigroup.

This is a major result since Baer*-semigroups and rings have a rich mathematical structure [3], and representation theory exists. For example compatible events correspond to commutativity of associated operations, as it should be. Classical systems correspond to *commutative* Baer* semigroups. Noncompatible events are properly represented in terms of operations; a fact extremely important for data fusion.

We then have

Theorem 4: Data bases and conditioning in a distributed asynchronous stochastic system can be used to construct an event-state-operation structure. This is a statistically faithful representation.

Starting from here we can develop structure theory, classification, and several very interesting connections with the theory of Baer*-rings, Rickart*-rings, C^* algebras. In particular we can show

Theorem 5: If $(E, \leq)$ is atomic and the atoms $\hat{E}$ are mapped to pure states under T, then (E, S, P, T) can be represented as the lattice of projections on a Hilbert space.

The assumption is typically valid in applications. Furthermore in applications the Hilber space is often finite dimensional.

Finally, we show how this structure can be embedded in a convex structure. This is important for optimization.

Starting from an event-state structure (E, S, P) one embeds S into the real vector space V of functions on E defined by

$$X(p) = \sum_{i=1}^{n} c_i\, P(p, a_i), \; a_i \in S$$

c_i real numbers, n arbitrary. We make V in a real Banach-space and consider V^*, its dual. Now one can identify E with a subset of V^*, and can introduce a partial order in V by a cone V^+. The states S are identified as the elements of $\{u \in V^+ : \tau(x) = 1\}$, where τ is the norm functional. The states form a convex set. We then have

Theorem 6: (E, S, P, T) can be mapped into a pair of Banach spaces V, V^* with positive cones and a trace functional τ. S is identified with a convex subset of V. E is the set of extreme points of a convex subset of V^*. Operations correspond to linear positive maps $T : V \longrightarrow V$ such that $0 \leq \tau(Tx) \leq \tau(x)$.

A generalized sensor on $(U, \mathcal{B})$ is a map $M : \mathcal{B} \longrightarrow \mathcal{L}^+(V)$ such that

$$M(B) \geq M(\phi)$$
$$M(\coprod B_i) = \Sigma\, M(B_i)$$
$$\tau(M(V)\rho) = \tau(p) \text{ forall } \rho \in V.$$

That is M is an operator valued measure. The interpretation is that a generalized sensor accepts a state, measures some properties and emits an output state conditioned on the value of the measurement (observation). Families of generalized sensors become important when we consider dynamic problems. For state problems we need only the *measurement* associated with a generalized sensor. Indeed for the problem of interest here, distributed detection or estimation, there is no action or control. The *mea-*

surement K_M associated with the generalized sensor M is the unique V^* valued measure such that

$$K_M(B)(\rho) = \tau[M(B)\rho$$
$$\forall \rho \in V, B \in \mathcal{B}.$$

Note that the statistics of the observed (or collected) data, when the system "state" is ρ, by a generalized sensor M are given by the probability measure $K_M(B)(\rho), \forall B \in \mathcal{B}$. We can consider the composition of two generalized sensors in time for example.

We have been able to interpret agent coordination, as necessary extension of certain constructs of the theory. Our theory recovers the observed statistics faithfully but cannot recover the actual measurements and communication strategies or histories employed by each agent or sensor.

We are now ready to describe our major results for the distributed M-arry detection problem. Suppose that the N agents (sensors) operate asynchronously over a time interval $[0, T]$. We collect all the local observation times t_k^i, $i = 1, \dots, N$ and globally order them. Each agent i, $i = 1, \dots, N$, at each local instant t_k^i, $i = 1, \dots, N$, $k = 1, \dots, L(i)$, has data $y^i(t_k^i)$ (his own) plus data $z^j(t_k^i)$, $j \neq i$, $j = 1, \dots, N$, communicated to him from other agents. z^j may be processed or unprocessed. We want to ask the following fundamental question. Given arbitrary communication, how can one represent the statistics of the collected data $(y^i(t_k^i), z^j(t_k^i))$, $i = 1, \dots, N$, $j \neq i, j = 1, \dots, N$, $k = 1, \dots, L(i)$? We have the following answer.

Theorem 6: In the distributed, M-ary detection problem described above, any sequence of observations and communications between the agents can be represented by an appropriate *measurement* $K_M(\cdot)$ on some measurable space $(U, \mathcal{B})$.

The proof is nonconstructive. This we consider as an important conceptual tool, particularly with respect to obtaining performance bounds. The latter is its greatest advantage. Its disadvantage, is that it is not possible to recover from K_M the actual observation process *and* the communication strategy.

In many situations, we have the Hilbert space model, and indeed a finite dimensional one. In such cases, one can perform numerical studies and obtain useful bounds with these methods. Then the bounds can be utilized to evaluate the performance of ad hoc communication strategies, for example.

Another important point that was made earlier is that in this specific setting,
Naimark's extension theorem provides a natural way of coordination between noncompatible observers. The fact that this comes out of the mathematical model automatically is a measure of success for the underlying models that we constructed.

To solve now the M-ary distributed detection problem in view of the representation result presented in Theorem 6, one proceeds as follows. Here, we concentrate on the Hilbert space model, but it should be clear by now how to extend the computation to more general models.

Given the M-hypotheses $H_1, \dots, H_M$, one constructs risk operators $W_i \in J_s(\mathcal{H})$, $i = 1, \dots, M$, based on assumed costs, states corresponding to the M hypotheses and prior probabilities. We allow, of course, randomized strategies, and we search for the optimal *measurement* $K_M(\cdot)$, subject to some information pattern constraint. Let

$$\Pi_i = \int_U \Pi_i(u) K_M(du), \quad i = 1, \dots, M$$

The problem becomes

$$\min T_r \sum_{i=1}^{M} W_i \Pi_i$$

over all positive operator valued measures (POM) Π_i, $i = 1, \dots, M$ such that

$$\{\Pi_i\}_{i=1}^M \in \mathcal{A}$$

Here $\mathcal{A}$ is a convex set of POM's corresponding to some *information theoretic* constraint on information (communication) patterns, such as capacity constraints, for example.

This problem is a convex linear programming problem and its duality theory is well understood (see, for example, [4]). One can then in this example begin to understand how the *duality* between decisions and information patterns can be put in a firm framework. Further work is needed along this promising direction, however. For example, for the unconstrained problem, we have the following result.

Theorem 7: Suppose $\mathcal{A}$ above is the set of all POM's. Then a necessary and sufficient condition for the POM Π_i^*, $i = 1, \dots, M$, to be optimal is that

$$(i) \quad \sum_{j=1}^{M} W_j \Pi_j^* \leq W_i, i = 1, \dots, M$$

$$(ii) \quad \sum_{j=1}^{M} \Pi_j^* W_j \leq W_i, i = 1, \dots, M$$

Furthermore, under any of the above conditions the operator

$$Y = \sum_{j=1}^{M} W_j \Pi_j^* = \sum_{j=1}^{M} \Pi_j^* W_j$$

is self-adjoint and is the unique solution of the dual problem.

It is easy to see that the above conditions are equivalent to Y being self-adjoint and

$$W_i \geq Y, \; i = 1, \dots, M.$$

Then these imply

$$(W_i - Y)\Pi_i^* = \Pi_i^*(W_i - Y); \; i = 1, 2, \dots, M$$

and that the minimum value is TrY.

We would like to close this section by mentioning that these results can be extended to include estimation problems. The major outstanding open problem is that of *implementation*. That is, if we find the optimal Π_i^*, how do we realize it by a communication pattern and a classical measurement process? It is also possible to interpret the Lagrange multipliers (here the Y), as sensitivities with respect to the information pattern constraints.

References

[1] J. S. Baras, "Noncommutative Probability Models in Quantum Communication and Multi-Agent Stochastic Control," in *Ricerche di Automation*, December 1979, Vol. 10, No. 2, pp. 217-264.

[2] J.S. Baras, *Distributed Detection with Asynchronous Computations*, submitted for publication.

[3] D.J. Foulis, *Baer*-Semigroups*, Proc. of Am. Math. Soc., 1960.

[4] S. Young, "Convexity and Duality in Optimization Theory," Ph.D. Dissertation, Mathematics Department, MIT, September 1977.