Cooperation and Optimization Strategies in Sequential Decision Making

Francesco Bullo

Center for Control,
Dynamical Systems & Computation
University of California at Santa Barbara
http://motion.me.ucsb.edu

MURI FA95500710528 Project Review: Behavioral Dynamics in Cooperative Control of Mixed Human/Robot Teams
Center for Human and Robot Decision Dynamics, Nov 12, 2009

Personnel at UCSB

Professor Francesco Bullo

Supported (including partially supported) personnel since project start:

1. Ruggero Carli, postdoc
2. Kurt Plarre, postdoc, now at University of Memphis
3. Sandra H. Dandach, PhD student (advanced to candidacy, dec 08)
4. Vaibhav Srivastava, PhD student (candidacy exam, dec 09)
5. Nathan P. Owen, MS '09, now at Boeing Space & Intelligence Systems

Awards and recognitions

Carli: 2009 EECI Best PhD Thesis, Best Student Paper Award @ 2009 IFAC NecSys.
Bullo: 2008 IEEE CSM Best Paper Award, Promotion to Professor in 2008, Plenary speaker at 2009 IEEE MSC, Keynote speaker at 2008 Int Conf Appl Math & Computing.

Selected Publications

Summary: journal articles (3 in print, 1 submitted), conference 8, chapters 1, book 1

1. V. Srivastava, K. Plarre, and F. Bullo. Randomized sensor selection in sequential hypothesis testing. IEEE Trans Signal Processing, September 2009. Submitted
2. S. H. Dandach, R. Carli, and F. Bullo. Accuracy and decision time for cooperative implementations of the sequential probability ratio test. In Proc ACC, Baltimore, MD, June 2010. Submitted
3. S. H. Dandach and F. Bullo. Algorithms for regional source localization. In *Proc ACC*, pages 5440–5445, St. Louis, MO, June 2009
4. K. Plarre and F. Bullo. Increasingly correct message passing averaging algorithms. In *Proc CDC*, pages 1304–1310, Cancún, México, December 2008
5. R. Carli, F. Bullo, and S. Zampieri. Quantized average consensus via dynamic coding/decoding schemes. *Int. J. Robust and Nonlinear Control*, 2009. (Submitted May 2008) to appear
6. F. Bullo, J. Cortés, and S. Martínez. *Distributed Control of Robotic Networks*. Applied Mathematics Series. Princeton Univ Press, 2009. Available at http://www.coordinationbook.info
7. M. Pavone, E. Frazzoli, and F. Bullo. Distributed and adaptive algorithms for vehicle routing in a stochastic and dynamic environment. *IEEE Trans Automatic Ctrl*, August 2009. (Submitted, Apr 2009) to appear
8. S. L. Smith, M. Pavone, F. Bullo, and E. Frazzoli. Dynamic vehicle routing with priority classes of stochastic demands. *SIAM JCO*, 2009. (submitted Feb 2009) to appear

Interactions and Collaborations

Interations with MURI members

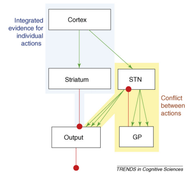
- Princeton: discussions about SPRT, human and collective decision making
- University of Washington: discussions and reciprocal visits
- Boston University: discussions about mixed networks and task allocation

Interactions with scientists outside the MURI

- Prof. Jeff Moehlis (ME and Math, UCSB): cooperative decision making, SPRT and human decision making
- Research Scientist Ketan Savla and Prof. Emilio Frazzoli (Aero, MIT): vehicle routing and human-in-the-loop designs
- Senior Engineer Anurag Ganguli (UtopiaCompression Inc): variable autonomy in mixed networks

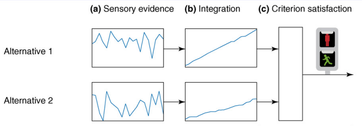
Human Decision Making Models

- decision process involves three steps
  1. evidence collection
  2. integration
  3. criterion satisfaction
- a central switch arbitrates decisions.
- sequential hypothesis testing models human decision making reasonably accurately



Brain activity during decision making (Picture from: R. Bogacz. "Optimal decision-making theories: linking neurobiology with behaviour" *Trends in Cognitive Sciences*, 11(3):118–125, 2007)

Sequential Hypothesis Testing



- sensory evidence ∼ likelihood ratio
- integration ∼ posterior probability
- criterion satisfaction ∼ crossing some threshold

SPRT = optimal decision time at given accuracy

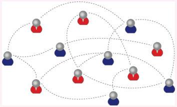
Outline: Decision Making in Networks

Topic #1: Cooperation and communication strategies

- How to aggregate individual decisions into network decisions?
- What accuracy/time tradeoffs characterize prototypical aggregation?

Topic #2: Sensor selection and attention control strategies

- How to identify most informative sensors?
- How to minimize decision time, at given accuracy level?



Cooperative Decision Making

Sensor Selection in Decision Making

Topic #1: Cooperative decision making

- Each equally-informed individual makes local decisions
- Network decision rule aggregates individual decisions

Model #1: network decision rule

How to integrate decisions by N nodes into collective decision:

1. **Leader rule:** pre-selected node decides for network
2. **Fastest rule:** fastest node decides for network
3. **Majority rule:** network agrees with majority decision

Model #2: communication protocols

1. **No-communication protocol** (independent decision makers)
2. **Sequential message passing protocol**

Topic #1: Cooperative decision making

Goal: examine accuracy & decision time as function of (comm-protocol × decision-rule)

	Leader	Fastest	Majority
No-communication	SPRT	?	?
Seq. message passing	?	?	?

accuracy and decision time in a network?

Sequential decision making: SPRT algorithm

- two alternative hypotheses: H_0, H_1
- repeated measurements $y_0, y_1, \dots$
  conditional probabilities $\mathbb{P}(y|H_0)$ and $\mathbb{P}(y|H_1)$ are known

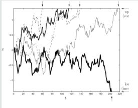
Sequential Probability Ratio Test

Given thresholds $\eta_0 < 0 < \eta_1$, at each time t

(a) log likelihood ratio $\lambda(t) := \log \frac{\mathbb{P}(y_t|H_1)}{\mathbb{P}(y_t|H_0)}$

(b) $\Lambda(t) := \sum_{\tau=1}^{t} \lambda(\tau)$

(c)
$$\eta_1 < \Lambda(t) \Rightarrow \text{say } H_1$$
$$\Lambda(t) < \eta_0 \Rightarrow \text{say } H_0$$
$$\eta_0 < \Lambda(t) < \eta_1 \Rightarrow \text{sample again}$$

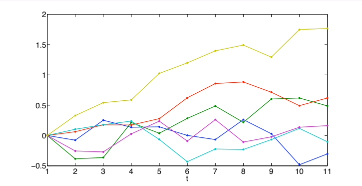


Quickest decision at given accuracy levels (A. Wald 1945)

$$\eta_0 = (1 - P_{\text{md}})/P_{\text{fa}}, \qquad \eta_1 = (1 - P_{\text{fa}})/P_{\text{md}}$$

No-communication SPRT

- each decision maker run individually a SPRT algorithm
- all the decision makers have the same thresholds η_0, η_1



(No-comm SPRT) + (Fastest rule)

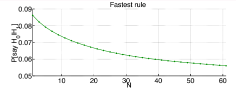
Theoretical results

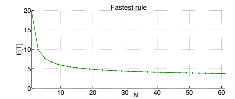
Accuracy

- $\mathbb{P}[\text{say } H_0|H_1, N+1] \leq \mathbb{P}[\text{say } H_0|H_1, N]$ (Monotonicity)

Expected Time

- $E[T_{\text{Fastest}}] \leq \mathbb{E}[T_{\text{SPRT}}]$
- $E[T_{\text{Fastest}}] \geq \frac{1}{\sqrt{N}} E[T_{\text{SPRT}}]$





(No-comm SPRT) + (Majority rule)

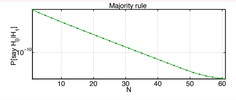
Theoretical results

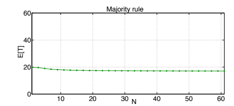
Accuracy

- $\mathbb{P}\,[\text{say } H_0|H_1, N+1] \leq \mathbb{P}\,[\text{say } H_0|H_1, N]$ (Monotonicity)
- $\mathbb{P}\,[\text{say } H_0|H_1, N] = \binom{N}{\lceil N/2 \rceil} p_f^{\lceil N/2 \rceil} + o\left(p_f^{\lceil N/2 \rceil}\right)$

Expected Time

- $E[T_{\text{Majority}}] \approx E[T_{\text{SPRT}}]$





Analysis 1/3

For a single decision maker, it is known how to compute

$$p_{i|j}(t) := \text{Probability "say } H_i \text{ given } H_j\text{" at time } t$$

$$\mathbb{P}[\text{say } H_i|H_j] = \sum_{t=1}^{+\infty} p_{i|j}(t), \qquad E[T_{\text{SPRT}}] = \sum_{t=1}^{+\infty} t\left(p_{1|i}(t) + p_{0|i}(t)\right)$$

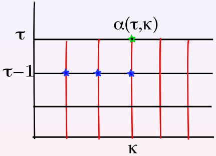
- discrete exponential family distributions (Young '94)
- continuous distributions approximated by absorbing Markov chain

Towards $p_{i|j}(t, N)$: intermediate probabilities

$\alpha(t,k)$	probability balanced events up to time t, $2k$ events
$\beta_{i|j}(t,k)$	probability decision (say H_i, given H_j) of k sensor at time t
$\gamma_{i|j}(t)$	probability (say H_i, given H_j) up to time t

Analysis 2/3

- $\alpha(t,k)$ is built using values at previous time $t-1$ up to k counts



- $\beta_{i|j}(t,k)$, $\gamma_{i|j}(t)$ computed as a function of $\{p_{1|j}(t), p_{0|j}(t)\}_{t=1}^{\infty}$
  - $\gamma_{i|j}(t) = p_{i|j}(1) + \ldots + p_{i|j}(t)$
  - $\beta_{i|j}(t) = f\left(p_{1|j}(t), p_{0|j}(t), \gamma_{1|j}(t), \gamma_{0|j}(t), \gamma_{1|j}(t-1), \gamma_{0|j}(t-1)\right)$

Analysis 3/3

Fastest rule

Probability that fastest rule gives correct decision at time t is

$$p_{i|j}(t, N) = \sum_{s=0}^{\lfloor \frac{N}{2} \rfloor} \binom{N}{2s} \alpha(t-1, s) \beta_{i|j}(t, N-2s)$$

Majority rule (N odd)

Probability that fastest rule gives correct decision at time t is

$$p_{i|j}(t, N) = \sum_{s=0}^{\lceil \frac{N}{2} \rceil - 1} \sum_{r=\lceil \frac{N}{2} \rceil - s}^{N-s} \binom{N}{s} \gamma_{i|j}^{s}(t-1) \times \binom{N-s}{r} p_{i|j}^{r}(t) \left(1 - \gamma_{i|j}(t)\right)^{N-(r+s)}$$

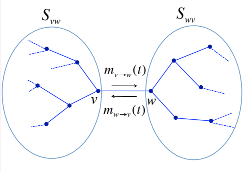
Cooperative decision making with communication

- $\mathcal{G} = (\mathcal{V}, \mathcal{E})$: undirected tree
- v-th node has a value x_v

Message passing algorithm for computing $\sum_{v \in \mathcal{V}} x_v$

$m_{v \to w}(t)$: estimate at time t of the sum of elements in S_{vw}

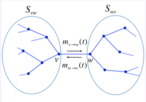
$m_{w \to v}(t)$: estimate at time t of the sum of elements in S_{wv}



At time t, node v knows sum of all values at distance $\leq t$

Cooperative decision making with communication

- $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ undirected tree
- $\mathcal{N}_v$ set of neighbors of node v
- v-th node memory: $m_{v \to w}$, $m_{w \to v}$, Λ_v



Sequential Message Passing SPRT (sequential evidence integration)

At each time t, node v

(i) take new measurement y_t and compute $\lambda(t) := \log \frac{\mathbb{P}(y_t|H_1)}{\mathbb{P}(y_t|H_0)}$

(ii) updates $\Lambda_v(t) = \sum_{\tau=1}^{t} \lambda_v(\tau) + \sum_{u \in \mathcal{N}_v} m_u(t-1)$

(iii) updates $m_{v \to w}(t) = \sum_{\tau=1}^{t} \lambda_v(\tau) + \sum_{u \in \mathcal{N}_v \setminus w} m_{u \to v}(t-1)$

(iv) transmits to $w \in \mathcal{N}_v$ the symbol $m_{v \to w}(t)$

(v) compares $\Lambda_v(t)$ to η_0 and η_1

(Sequential MP) + (Leader rule)

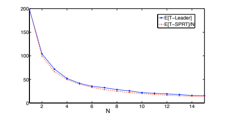
Theoretical results

Accuracy

- $\mathbb{P}\,[\text{say } H_1 | H_0] \leq P_{\text{fa}}$
- $\mathbb{P}\,[\text{say } H_0 | H_1] \leq P_{\text{md}}$

Expected Time

- $\mathbb{E}\,[T_{Leader}] \leq \frac{\mathbb{E}[T_{SPRT}]}{N} + diam$



(Sequential MP) + (Fastest rule)

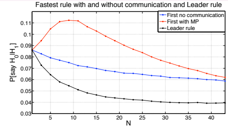
Monte-Carlo simulations

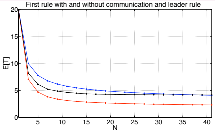
Accuracy

- No-monotonicity
- worsens with communication

Expected time

- improves with communication





(Sequential MP) + (Majority rule)

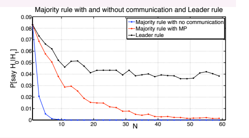
Monte-Carlo simulations

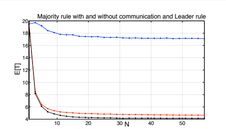
Accuracy

- monotonically decreases with number of agents
- worsens with communication

Expected time

- improves with communication





Summary of results

rescale accuracy of local decisions so that network accuracy is equal

	Leader	Fastest	Majority
No-communication	unit time	$> \frac{1}{\sqrt{N}}$	Better than Fastest + No-Comm
Seq. message passing	$\frac{1}{N} + diam$	$\frac{1}{N} + diam$	$\frac{1}{N}$

Topic #1: Conclusions and Future Work

Conclusions

- Comprehensive analysis of multiple scenarios
- Results rely on Wald's classic approximation
- Rigorous analysis of no-communication scenarios, and MonteCarlo analysis of (fastest or majority) + (sequential MP)

Future research directions

1. Optimal strategies for a desired accuracy in group decision
2. Interactions of different groups with various decision rules
3. Cooperative SPRT with social communications/interactions

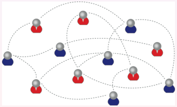
Outline: Decision Making in Networks

Topic #1: Cooperation and communication strategies

- How to aggregate individual decisions into network decisions?
- What accuracy/time tradeoffs characterize prototypical aggregation?

Topic #2: Sensor selection and attention control strategies

- How to identify most informative sensors?
- How to minimize decision time, at given accuracy level?



Cooperative Decision Making

Sensor Selection in Decision Making

Topic #2: Sensor selection and attention control strategies

Attention in Camera Sensor Network

Which camera to choose?

Sensors for UAV Surveillance

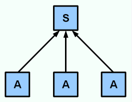
	Cloud	Rain	Wind
EO	X	X	X
SAR		X	
FPR		X	X
IR	X	X	
GMTI		X	

Which sensor to choose?

1. how to avoid operator overload
2. how to select most informative sensors and focus attention

Topic #2: Decision making with switching sensors

- n sensors transmit their observations to a fusion center, one at a time
- processing time of sensor s is T_s
- randomized sensor selection strategy
- sensor selection metric: decision time



SPRT with switching sensors

1. compute log likelihood ratio:

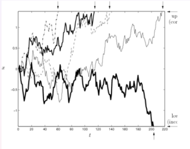
$$\lambda_l := \log \frac{f^1_{s_l}(y_{s_l})}{f^0_{s_l}(y_{s_l})},$$

2. integrate evidence up to time N, i.e.,

$$\Lambda_N = \sum_{l=0}^{N} \lambda_l,$$

3. decide only if a threshold is crossed, i.e.,

$$\begin{cases} \Lambda_N > \eta_1, & \text{say } H_1, \\ \Lambda_N < \eta_0, & \text{say } H_0, \\ \Lambda_N \in]\eta_0, \eta_1[, & \text{continue sampling} \end{cases}$$



SPRT evolutions

MSPRT with switching sensors

1. define likelihood ratios

$$\lambda^{k,j}_l = \frac{f^j_{s_l}(y_l)}{f^k_{s_l}(y_l)}, \quad \forall j \in \{0, \dots, M-1\} \setminus \{k\}.$$

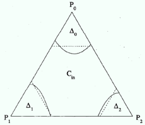
2. integrate the evidence

$$\Lambda^k_N = \sum_{\substack{j=0 \\ j \neq k}}^{M-1} \prod_{l=1}^{N} \lambda^{k,j}_l.$$

3. decide on a hypothesis, if some threshold is crossed

$$\begin{cases} \Lambda^k_N < \eta_k, \text{for at least one } k, & \text{say } H_{\bar{k}}, \\ \text{otherwise}, & \text{continue sampling} \end{cases}$$

Condition of audience?



MSPRT decision regions

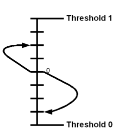
Expected decision time

Expected decision time

Expected decision time of SPRT/MSPRT conditioned on hypothesis H_k is

$$\mathbb{E}[T_d|H_k] = \frac{q \cdot T}{q \cdot I^k}, \qquad q \in \Delta_{n-1}, T, I^k \in \mathbb{R}_{>0}.$$

Idea of the proof:



expected decision time

$$= \frac{\text{threshold}}{\text{expected jump}} \times \text{expected processing time.}$$

Optimal sensor selection

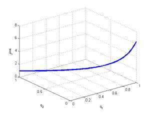
Optimization of conditioned decision

Minimize: $J_k^{\text{one}}(q) = \mathbb{E}_q[T_\text{d}|H_k]$.

The **optimal probability vector** is: $q^* = e_{s^*}$, $s^* = \underset{s \in \{1,\dots,n\}}{\operatorname{argmin}} \frac{T_s}{I_s^k}$.

Idea of proof:

- $\mathbb{E}_q[T_\text{d}|H_k]$ is a **linear-fractional function**
- a linear-fractional function is monotone
- minima/maxima of a linear-fractional function lies at an extreme point of the feasible region



Optimal strategy for change detection

Optimal sensor selection

Worst case optimization

Minimize: $J^{\max}(q) = \max\{\mathbb{E}_q[T_\text{d}|H_0], \mathbb{E}_q[T_\text{d}|H_1]\}$

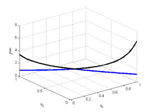
optimal probability vector is: $q^* = \begin{cases} e_{w^*}, & \text{if } J^{\max}(w^*) \leq J^{\max}(s^*, r^*), \\ q^{s^* r^*}, & \text{if } J^{\max}(w^*) > J^{\max}(s^*, r^*). \end{cases}$

$$(s^*, r^*) \in \underset{\substack{I_s^0 - I_s^1 \leq 0 < I_r^0 - I_r^1 \\ s,r \in \{1,\dots,n\}}}{\operatorname{argmin}} \{J^{\max}(s,r)\}, \text{ and } w^* = \underset{w \in \{1,\dots,n\}}{\operatorname{argmin}} \{J^{\max}(e_w)\}$$

Idea of proof:

- minima lie at the boundary or the intersection of graphs

at most two sensors are optimal



Optimal sensor selection

Average decision time optimization

Minimize: $J^{\text{avg}}(q) = \frac{1}{M}\left(\mathbb{E}_q[T_\text{d}|H_0] + \dots + \mathbb{E}_q[T_\text{d}|H_{M-1}]\right)$.

For a generic set of sensors,

1. at most M sensors need to be observed.
2. and $M = 2$, **optimal probability vector** is $q^* = \underset{s,r \in \{1,\dots,n\}}{\operatorname{argmin}} \{J^{\text{avg}}(q^*_{sr})\}$.

Idea of proof:

- $T, I^0, \dots, I^{M-1}$ are linearly independent
- $J^{\text{avg}}(q)$ is monotone for $n > 2$
- minimum lies at an edge of the simplex

Topic #2: Conclusions and Future Work

Conclusions

- Bottom line: focus on few key sensors
  (same number as number of hypotheses)
- Complete results for SPRT, partial results or MSPRT
- Simple model leads to nontrivial results
- Results rely on Wald's classic approximation

Future research directions

1. Learning scheme for probabilities
2. More general sensor selection strategies: Markov decision processes
   (connections with human decision making)

Cooperation and Optimization in Decision Making

Scientific foundations

Stochastic decision making
Algorithmic coordination
Convex and linear-fractional optimization

Accomplishments

Efficient cooperative SPRT designs
Optimal sensor selection for decision making
Application of distributed decision making to source localization

Ongoing work

1. Cooperative decision making: comparisons + analysis of MP
2. Sensor selection: learning schemes + more general strategies
3. Integration with vehicle routing scenarios
4. Applications to social networks (models of social decision making)