O Tempora, O Mores

Bill Moran

Melbourne Systems Laboratory

March 5, 2007

Overview

- Liverpool and Measure Algebras
- Cambridge, Normal Numbers and Schmidt's Conjecture
- Sensor Scheduling
- The OPC Conjecture

In the Beginning

- $\mathcal{S}$ — compact single generator (monothetic) semigroup. $(r,s) \to r.s$ continuous (joint continuity) $\implies \exists!$ idempotent. This is the unit for a compact group $\mathcal{K}$ — *kernel* of the semigroup.
- What happens if $r \to r.s_0$ and $s \to r_0.s$ are continuous for all $r_0, s_0 \in \mathcal{S}$ (separate continuity)? Trevor West showed there can be more than one idempotent.

- (B & M 1971): The idempotent subsemigroup of a compact separately continuous monothetic semigroup can be an *arbitrary* lower semilattice.

Measure Algebras

- Algebra of measures $M(\mathbf{T})$ on circle $\mathbf{T} = \{e^{2\pi it} : t \in [0,1)\}$
- Look at complex homomorphisms $\Delta = \{\chi : M(\mathbf{T}) \to \mathbf{C} : \chi \text{ homomorphism}\}$.
- $\chi \in \Delta$ corresponds to $(\chi_\mu)_{\mu \in M(\mathbf{T})}$ where $\chi_\mu \in L^\infty(\mu)$:
  $\Delta_\mu = \{\chi_\mu : \chi \in \Delta\}$
- West used measure on Kronecker set K:

$$D \triangleq \mathrm{cl}\{e^{2\pi int} : n \in \mathbf{Z}\} = \text{ unit ball of } C(K)$$

  In general $D \subset \Delta_\mu$
- (Joe Taylor, Barry Johnson) $\exists$ singular measures such that $\chi_\mu(t) = ae^{2\pi int}$ for some $n \in \mathbf{Z}$, $a \in \mathbf{C}$, $\forall \chi \in \Delta$ (*tameness*)

Infinite Convolutions

- Bernoulli convolutions: $\mu = \bigstar_{n=1}^{\infty} \frac{1}{2}\big(\delta(-a_n) + \delta(a_n)\big)$
- More generally:

$$\mu = \bigstar_{n=1}^{\infty} \Big(\sum_{k,n} a_{k,n}\delta(x_{k,n})\Big) \tag{1}$$

- (B& M) Many Bernoulli convolutions are tame — arithmetical constraints on a_n's.
- Leads to monothetic semigroup result
- (B & M) Structure of Δ_μ for Bernoulli convolutions
- *Monotrochic:* $|\chi_\mu|$ constant for all $\chi \in \Delta$
- (B & M) Measures of form (1) are monotrochic
- (B & M) μ of form (1) implies one of following is true:
  - μ is discrete
  - $\mu^n \in L^1(\mathbf{T})$ for some n
  - $\mu^n \perp \mu^m$ for $n \neq m$

Orsay and Indiana

Brown

- Silov boundary is a proper subset of Δ_0 — maximal ideal space of $M_0(\mathbf{T})$
- $d\mu(t) = \prod_n (1 + a_n \cos 2\pi(r_n t + \phi_n)) dm(t)$ $(r_{n+1}/r_n > 3, a_n \geq 0)$
- Riesz products are tame, etc

Moran

- Silov boundary is a proper subset of Δ_0
- $F : \{z : |z| \leq 1\} \to \mathbf{C}$ continuous & $F(\hat{\mu}(n)) = \hat{\nu}(n) \forall n$. What does this say about F?
- If μ on Kronecker set then F analytic, etc

Orthogonality of Riesz Products

- Let

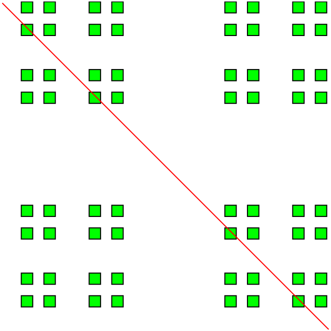
$$d\mu(t) = \prod_n (1 + a_n \cos(2\pi r_n t + \phi_n)).dm(t)$$
$$d\nu(t) = \prod_n (1 + b_n \cos(2\pi r_n t + \psi_n)).dm(t)$$

- (Jacques Peyriére) If $\sum_n |a_n e^{2\pi i \phi_n} - b_n e^{2\pi i \psi_n}|^2 = \infty$ then $\mu \perp \nu$.
- (B & M) If $\sum_n \frac{|a_n e^{2\pi i \phi_n} - b_n e^{2\pi i \psi_n}|^2}{2 - |a_n e^{2\pi i \phi_n} + b_n e^{2\pi i \psi_n}|} < \infty$ then $\nu \sim \mu$.

M_0, Boundaries, and Gleason Parts

- $M_0(\mathbf{T})$: measures μ whose Fourier transform
$$\widehat{\mu}(n) = \int_{\mathbf{T}} e^{-2\pi i n t}\, d\mu(t) \tag{2}$$
vanishes at infinity.
- $\Delta_0 = \Delta(M_0)$ is an open subset of Δ
- *A Boundary* is a subset B of Δ such that for every $\mu \in M$ there exists $\phi \in B$
$$|\phi(\mu)| = \sup_{\psi \in \Delta} |\psi(\mu)| \tag{3}$$
- (B & M) All boundaries for M_0 are boundaries for M
- (B & M) Characterise Gleason parts of measure algebras (Miller's Conjecture)

Measures on Cantor Sets and Woodall's Inequality



- *Lebesgue:* Let C be the classical ("middle-third") Cantor set on $[0, 1]$. Then $C + C = [0, 2]$.
- *Conjecture (B & M)* If A is a set of positive Cantor measure (μ_c) then $A + A$ is a set of positive Lebesgue measure (m). Reduced it to:
- *Woodall:*

$$x^a y^a + \max\{x^a(1-y)^a, y^a(1-x)^a\} + (1-x)^a(1-y)^a \geq 1$$

$(0 \leq x, y \leq 1)$, $a = (\log 3)/(\log 4)$

- (B & M) $m(E + F) \geq 2\mu_c(E)^a \mu_c(F)^a$.

Normality and Riesz Products

- *Schmidt's Theorem*: m, n positive integers (> 1) then $\exists$ real numbers x s. t. x is normal in base m but not in base n provided $\nexists$ solution to $n^r = m^s$ in integers r, s
- Original proof of Schmidt: effectively find infinite convolution measure $\mu_{m,n}$ s.t. x is normal in base m but not in base n almost surely wrt $\mu_{m,n}$

Normality and Riesz Products

- *Schmidt's Theorem*: m, n positive integers (> 1) then $\exists$ real numbers x s. t. x is normal in base m but not in base n provided $\nexists$ solution to $n^r = m^s$ in integers r, s
- Original proof of Schmidt: effectively find infinite convolution measure $\mu_{m,n}$ s.t. x is normal in base m but not in base n almost surely wrt $\mu_{m,n}$
- (B, M, Charles Pearce) Construct Riesz product $\mu_{m,n}$

A Month in Cambridge

- *Schmidt's conjecture*: Let S, T be $r \times r$ rational matrices which are *ergodic* — ie almost all (wrt Lebesgue measure) $\mathbf{x} \in \mathbf{R}^r$ are normal wrt S and T.
- Can we find $\mathbf{x} \in \mathbf{R}^r$ normal in base T but not normal in base S?
- *Normal* means $T^n\mathbf{x}$ is uniformly distributed modulo 1 in each coordinate.

- (B& M) If S and T commute Schmidt's conjecture is true.
- (B& M) If S and T are 2×2 and have real eigenvalues Schmidt's conjecture is true.

Enter Andy

- (B,M, Andy Pollington) Schmidt's conjecture is true in 2 dimensions
- In 1 dimension, free n and m from being integers — just reals $\alpha, \beta > 1$.
- Let $B(\alpha)$ be all numbers x normal in base β — ie $\beta^n x$ uniformly distributed modulo 1.
- THEOREM (B,M, Pollington)
  1. $B(\beta^r) \subset B(\beta^s) \quad (r \neq s) \iff \exists K : \beta^K \in \mathbf{N}$ & $\mathbf{Q}(\beta^r) \subset \mathbf{Q}(\beta^s)$ or $\beta^K + \beta^{-K} \in \mathbf{N}$
  2. $B(\lambda) \subset B(\tau) \implies \exists \beta, r, s : \lambda = \beta^r, \tau = \beta^s$ & 1. above holds
  3. $B(\lambda) = B(\tau) \iff \mathbf{Q}(\lambda) = \mathbf{Q}(\tau)$, $\log \lambda / \log \tau \in \mathbf{Q}$, & $\exists K : \lambda^K \in \mathbf{N}$

And now for something completely different

General Problem

Several evolving systems viewed in different ways under our control. Knowledge of systems and measurements have uncertainty. How to schedule measurements to minimize uncertainty?

Simple Example

- R systems with linear dynamics:

$$\mathbf{x}_n^{(r)} = F\mathbf{x}_{n-1}^{(r)} + \mathbf{w}_n^{(r)}$$

$\mathbf{w}_n^{(r)}$ is gaussian, mean $\mathbf{0}$, covariance $\Sigma_{\mathbf{w}^{(r)}}$

- Linear measurements:

$$\mathbf{y}_n^{(r)} = H_k\mathbf{x}_n^{(r)} + \mathbf{v}_n^{(r,k)}$$

$\mathbf{v}_n^{(r,k)}$ gaussian noise, covariance $\Sigma_{\mathbf{v}^{(r,k)}}$

Gauss-Markov Systems

- Suppose only one system and one measurement — H_1: Kalman filter gives optimal solution: minimum variance unbiased estimator for state $\mathbf{x}_n$ at time n based on all measurements $\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_n$
- Suppose one system and K measurements: find choice $H_{\pi(1)}, H_{\pi(2)}, \ldots, H_{\pi(n)}$ at time n to minimize summed (traces or determinants) of covariances of Kalman estimators at all times up to n
- Can be done offline — do not need to know state since covariance of estimator is a function of covariances $\Sigma_{\mathbf{w}}$, $\Sigma_{\mathbf{v}^{(k)}}$, and F (Kalman)
- But how to do it?

Even Simpler Problem

- Two one dimensional systems — states $\mathbf{x}_n^{(r)}$ and measurements $\mathbf{y}_n^{(r,k)}$ $(r = 1, 2)$ are one dimensional, linear maps are scalars

Even Simpler Problem

- Two one dimensional systems — states $\mathbf{x}_n^{(r)}$ and measurements $\mathbf{y}_n^{(r,k)}$ $(r = 1, 2)$ are one dimensional, linear maps are scalars
- Only two kinds of measurements — one of which is not to measure.

Even Simpler Problem

- Two one dimensional systems — states $\mathbf{x}_n^{(r)}$ and measurements $\mathbf{y}_n^{(r,k)}$ $(r = 1, 2)$ are one dimensional, linear maps are scalars
- Only two kinds of measurements — one of which is not to measure.
- Only need to track variances of estimators

Even Simpler Problem

- Two one dimensional systems — states $\mathbf{x}_n^{(r)}$ and measurements $\mathbf{y}_n^{(r,k)}$ $(r = 1, 2)$ are one dimensional, linear maps are scalars
- Only two kinds of measurements — one of which is not to measure.
- Only need to track variances of estimators
- $u_n^{(1)}, u_n^{(2)}$ are variances for systems at time n.

Even Simpler Problem

- Two one dimensional systems — states $\mathbf{x}_n^{(r)}$ and measurements $\mathbf{y}_n^{(r,k)}$ $(r = 1, 2)$ are one dimensional, linear maps are scalars
- Only two kinds of measurements — one of which is not to measure.
- Only need to track variances of estimators
- $u_n^{(1)}, u_n^{(2)}$ are variances for systems at time n.
- Asssume systems have same process noise and measurements have same measurement variance for each system.

Even Simpler Problem

- Two one dimensional systems — states $\mathbf{x}_n^{(r)}$ and measurements $\mathbf{y}_n^{(r,k)}$ $(r = 1, 2)$ are one dimensional, linear maps are scalars
- Only two kinds of measurements — one of which is not to measure.
- Only need to track variances of estimators
- $u_n^{(1)}, u_n^{(2)}$ are variances for systems at time n.
- Asssume systems have same process noise and measurements have same measurement variance for each system.
- After some normalisations:

$$u_n^{(1)} = u_{n-1}^{(1)} + 1$$
$$u_n^{(2)} = \frac{u_{n-1}^{(2)} + 1}{cu_{n-1}^{(2)} + c + 1}$$

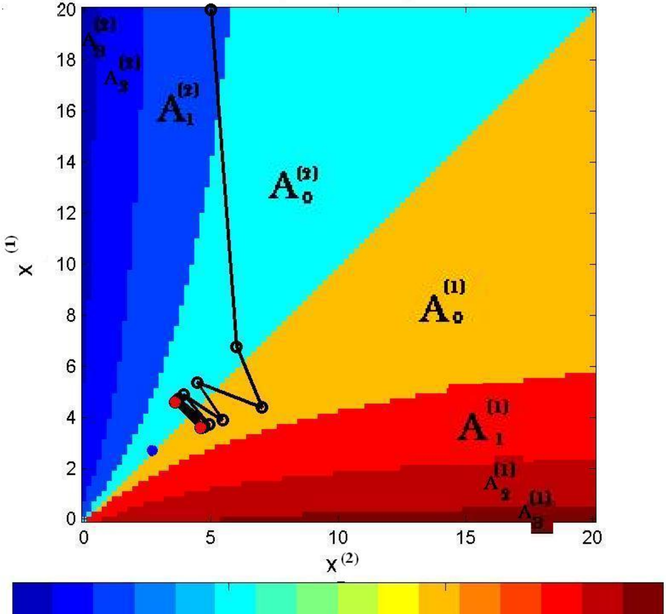
if we measure system 2, and roles reversed if we measure system 1.

- Cost function is

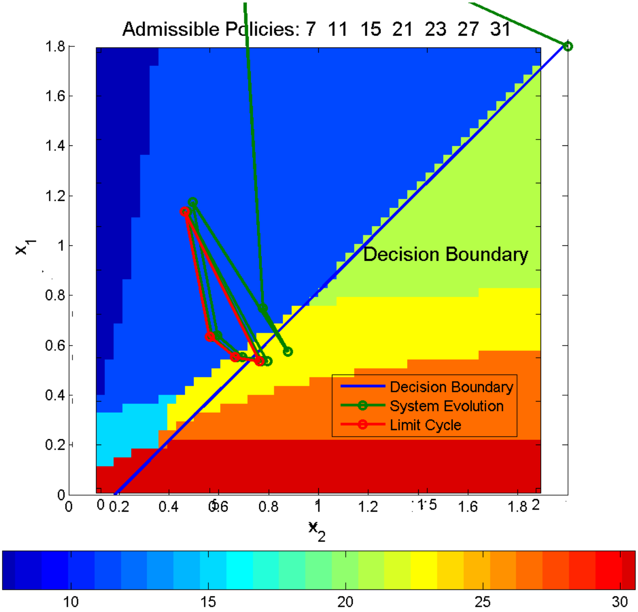
$$C_N(\mathbf{u}, \pi) = \sum_{n=1}^{N} u_n^{(1)} + u_n^{(2)}$$

  where π is a sequence in $\prod_{n=1}^{N}\{1,2\}$
- Find choice of π to minimise cost.

The solution



More Generally



Scheduling for HMMs

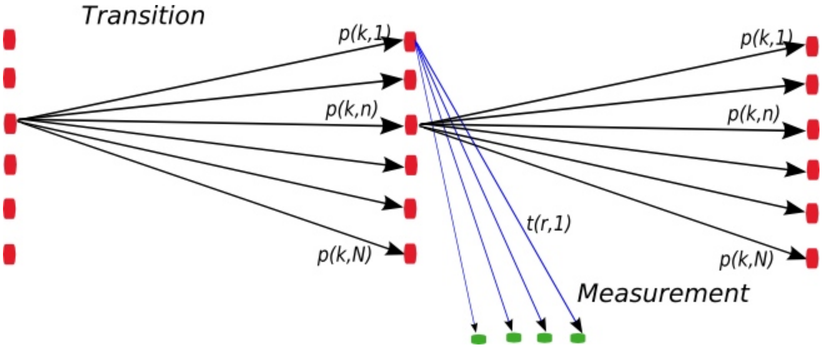
Hidden Markov Model

- $\mathcal{S}$ — state space — finite — size M
- P — stochastic transition matrix ($M \times M$)
- T — measurement matrix ($R \times M$)
- Δ — probability distributions on $\mathcal{S}$

Other definitions

- $\mathbf{S}_n$ — state at time n
- $\mathbf{Z}_n$ — measurement at time n
- $\mathcal{P}(\Delta)$ — probability measures on Δ

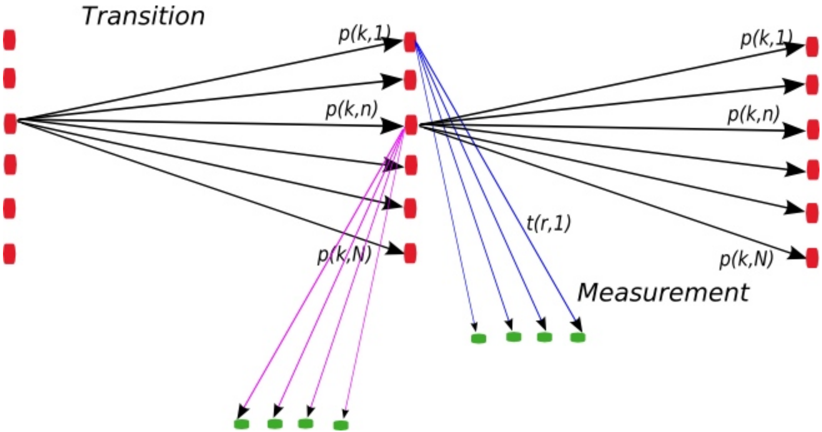
Hidden Markov Models



Multiple Measurements

- Different measurement matrices $\mathrm{T}^{(k)}$
- Cost function: Minimize uncertainty of next state of system given measurements: $H(\mathbf{S}_{n+1}|\mathbf{Z}^n)$
- Stationary: Make choice of measurement depend *only* on information state π_n — probability vector in Δ
- Can estimate information state from previous measurements — Bayes Rule update
- Find long term minimal cost — $\lim_n H(\mathbf{S}_{n+1}|\mathbf{Z}^n)$ based on a stationary policy

Hidden Markov Models



Description and Notation

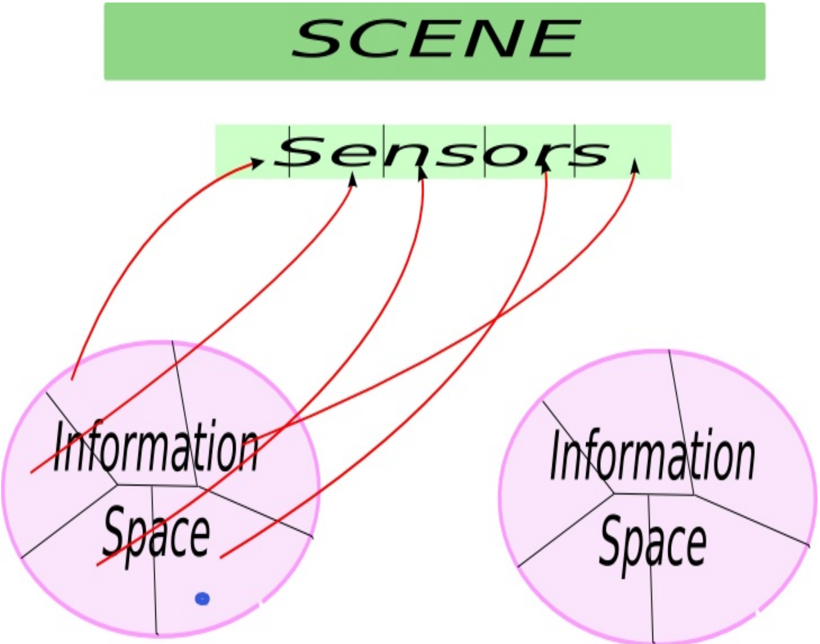
- π_n — posterior distribution of S_n at time n: $\pi_n = p(S_n|Z^{n-1})$;
- π_{n+1} — posterior distribution of S_n at time $n+1$:

$$\pi_{n+1} = p(S_n|Z^n) = f^{(k)}(z, \pi_n) = \frac{\pi_n D^{(k)}(z) P}{\pi_n D^{(k)}(z) \underline{1}},$$

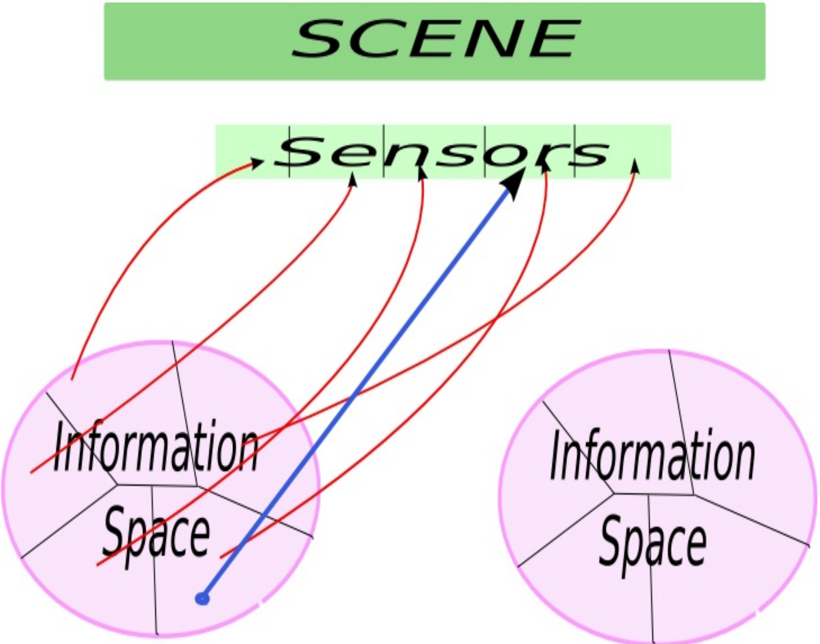
  where $D^{(k)}(z)$ — diagonal matrix with $d_{ii}(z) = T^{(k)}[i, z]$.
- Entropy rate for the state of the process:

$$\lim_{n\to\infty} H(S_n|Z^{n-1}) = \lim_{n\to\infty} \int_{\Delta} h(\pi_n) d\mu_n(\pi_n) = \int_{\Delta} h(\pi) d\mu(\pi).$$

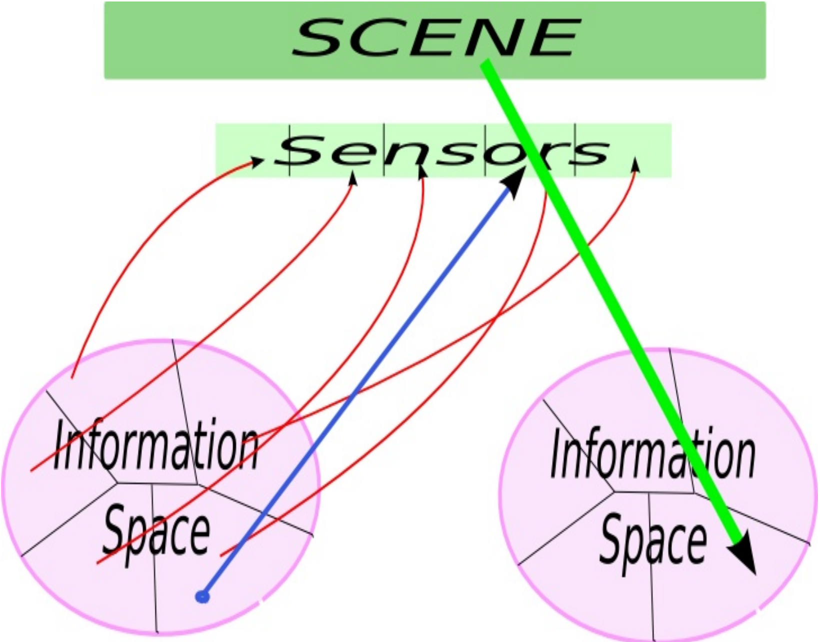
HMM Scheduling



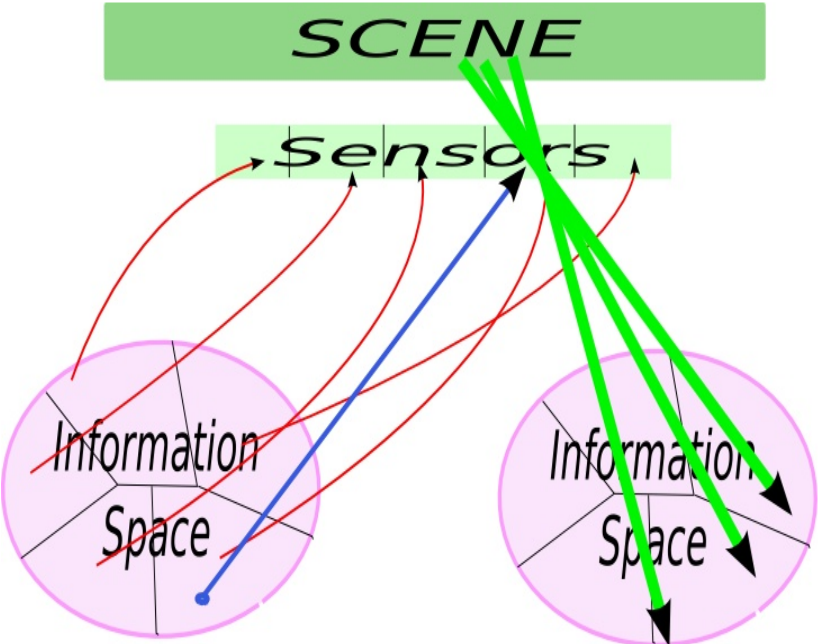
HMM Scheduling



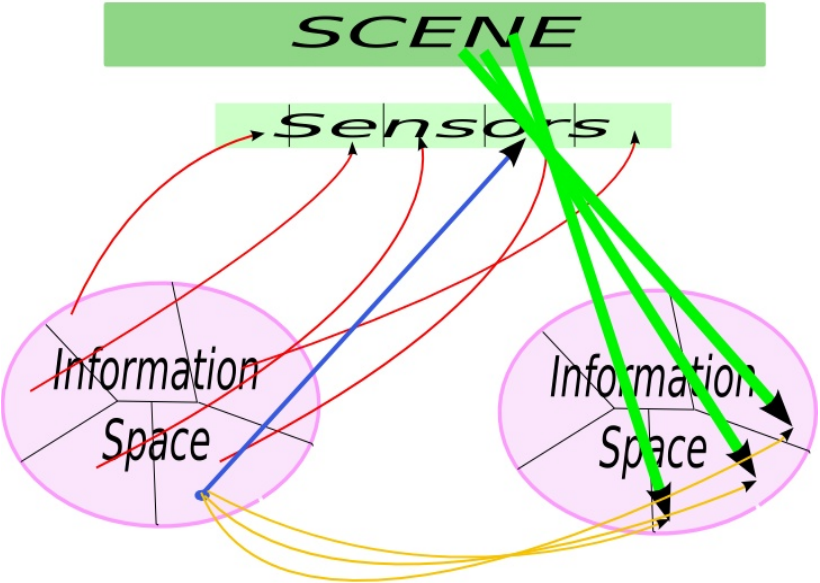
HMM Scheduling



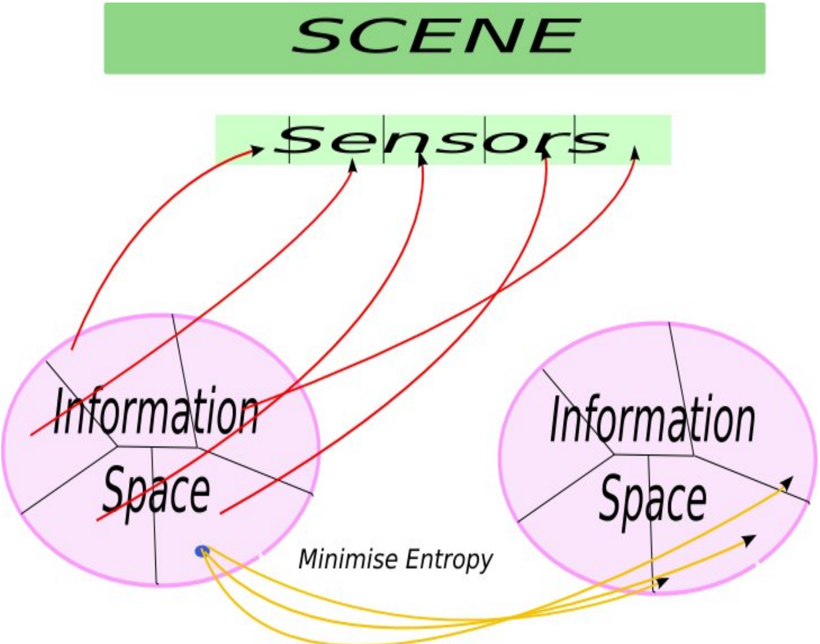
HMM Scheduling



HMM Scheduling



HMM Scheduling



Iterative Formula

- Distribution $\mu(.)$ obtained iteratively

$$\mu_{n+1}(\pi_{n+1}) = \int_{\Delta} \phi^{(k)}(\pi_n) d\pi_n$$

where

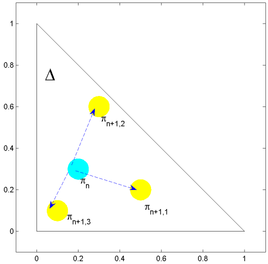
$$\phi^{(k)}(\pi) = \sum_{z} (\pi T^{(k)})_z \delta(f^{(k)}(z, \pi))$$

maps $\Delta \to \mathcal{P}(\Delta)$

- Starting from $\mu_0 = \delta(\nu), \pi_0 = \nu$

- Generate sets $\{\pi_n\}_i, \; i = 1, \ldots, |Z|^n$ and prob. distribution $\mu_n(\pi_n)$.
- Entropy Rate

$$H_n = \sum_{i=1}^{|Z|^n} \mu_n(\pi_{n,i}) h(\pi_{n,i}).$$



Stationary Policy

- *A stationary policy* is a partition $\tau = \{B_1, B_2, \dots B_M\}$ of the state space Δ by Borel sets; $\bigcup_{i=1}^{M} B_i = \Delta$.
- Define
$$\phi^{\tau}(\pi) = \sum_k \phi^{(k)}(\pi)\chi(B_k).$$
- Permits the definition of a map $\mathcal{P}(\Delta)$ to $\mathcal{P}(\Delta)$:
$$\Phi^{(\tau)}(\mu) = \int_{\Delta} \phi^{\tau}(\pi) d\mu(\pi) = \sum_k \int_{B_k} \phi^{(k)}(\pi) d\mu(\pi).$$

The Objective

- Find a policy $\tau^\star$ such that

$$H(\tau^\star) = \int_\Delta h(\pi) d\mu^{\tau^\star}(\pi) = \inf_\tau H(\tau).$$

Existence and Uniqueness of the Stationary Distribution

- Under suitable conditions on τ , $\Phi^{(\tau)}$ is a continuous convex map on the compact convex set $\mathcal{P}(\Delta)$ — has a fixed point:

$$\mu^{\tau}(\pi) = \Phi^{(\tau)}(\mu^{\tau}(\pi)).$$

- To form a fixed point

$$\rho_N^{\tau}(\mu) = \frac{1}{N+1}\sum_{n=0}^{N}(\Phi^{\tau})^n(\mu).$$

- Need to show independence of

$$\lim \rho_N^{\tau}(\delta(\pi))$$

from π to prove *uniqueness*.

Invariant Measure Lemma

The entropy rate of the state process is equivalent to:

$$H(\tau) = \int\limits_{\Delta} \underline{h}(\phi^{\tau}(\pi)) d\mu^{\tau}(\pi) = \int\limits_{\Delta} h(\pi) d\Phi^{(\tau)}(\mu^{\tau}(\pi)),$$

where $\underline{h}(\nu) = \int h d\nu$ for $\nu \in \mathcal{P}(\Delta)$.

The OPC Conjecture — Introduction

- Overflow loss networks: large and important class of loss networks (e.g. telephone networks).
- Exact performance solutions not scalable and only apply to cases where dimensionality is very small.
- Approximations required to estimate blocking probability

- Most used technique: Erlang Fixed-Point Approximation (1964)
- A new approximation called Overflow Priority Classification Approximation (OPCA) proposed (Zukerman *et. al.*) to improve EFPA.

The OPC Conjecture: Statement

- For simple and pure overflow loss network, numerical results show that the blocking estimated by OPCA (i.e. P_{OPCA}) lies between those estimated by the exact solution (i.e. P_{exact}) and by EFPA (i.e. P_{EFPA}):

$$P_{\text{exact}} \geq P_{\text{OPCA}} \geq P_{\text{EFPA}}$$

- Second inequality relatively easy to prove; first difficult — P_{OPCA} is a very good approximation to P_{exact}

The Gory Details

$$P_{\text{OPCA}} = 1 - \frac{\sum_{n=0}^{N-1} a(n)}{a\left[1 + \sum_{n=0}^{N-1} a(n)\right]} = \frac{(a-1)\sum_{n=0}^{N-1} a(n) + a}{a\left[1 + \sum_{n=0}^{N-1} a(n)\right]}$$

where

$$a(n) = \frac{\left[\sum_{i=0}^{n-1} a(i)\right]^2}{1 + \sum_{i=0}^{n-1} a(i)} - \sum_{i=1}^{n-1} a(i) = \frac{(a-1)\sum_{i=0}^{n-1} a(i) + a}{1 + \sum_{i=0}^{n-1} a(i)}$$

and $a(0) = a$. The blocking probability for the Erlang B exact solution

$$P_{\text{exact}} = \frac{\frac{(Na)^N}{N!}}{\sum_{n=0}^{N} \frac{(Na)^n}{n!}}$$

Theorem

(M, Wong, Zalesky, Zukerman)

$$P_{exact} \geq P_{OPCA} \qquad \forall N$$