IEOR 4615: Service Engineering
Lecture 8, February 17, 2015

- **Queueing Network Models** (Tuesday)
  1. **Classical Markov OQN Model** (Sec. 7 of CTMC notes)
     1. Yet another CTMC, but multi-dimensional
     2. Product-Form Distribution
     3. Traffic Rate Equations
  2. **non-Markov OQN: Approximations** (QNA,1983, 1989)
     1. Quantifying Variability and its Impact approximately
     2. The Parametric-Decomposition Approximation
     3. From Customer Routes to Markovian Routing
- Processing Networks (last Thursday&Friday)
  - Precedence Constraints , PERT/CPM

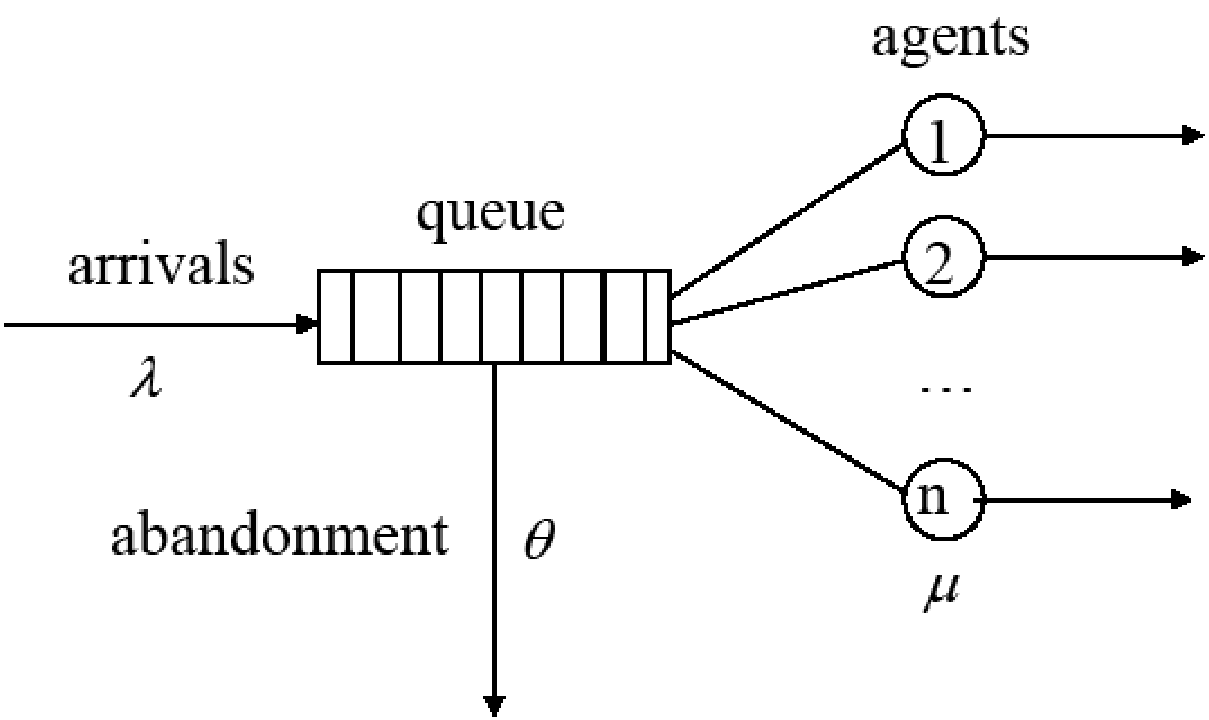
What do you see?

Need to have expectations

Experience and Models Help Greatly

Before
A Large Call Center

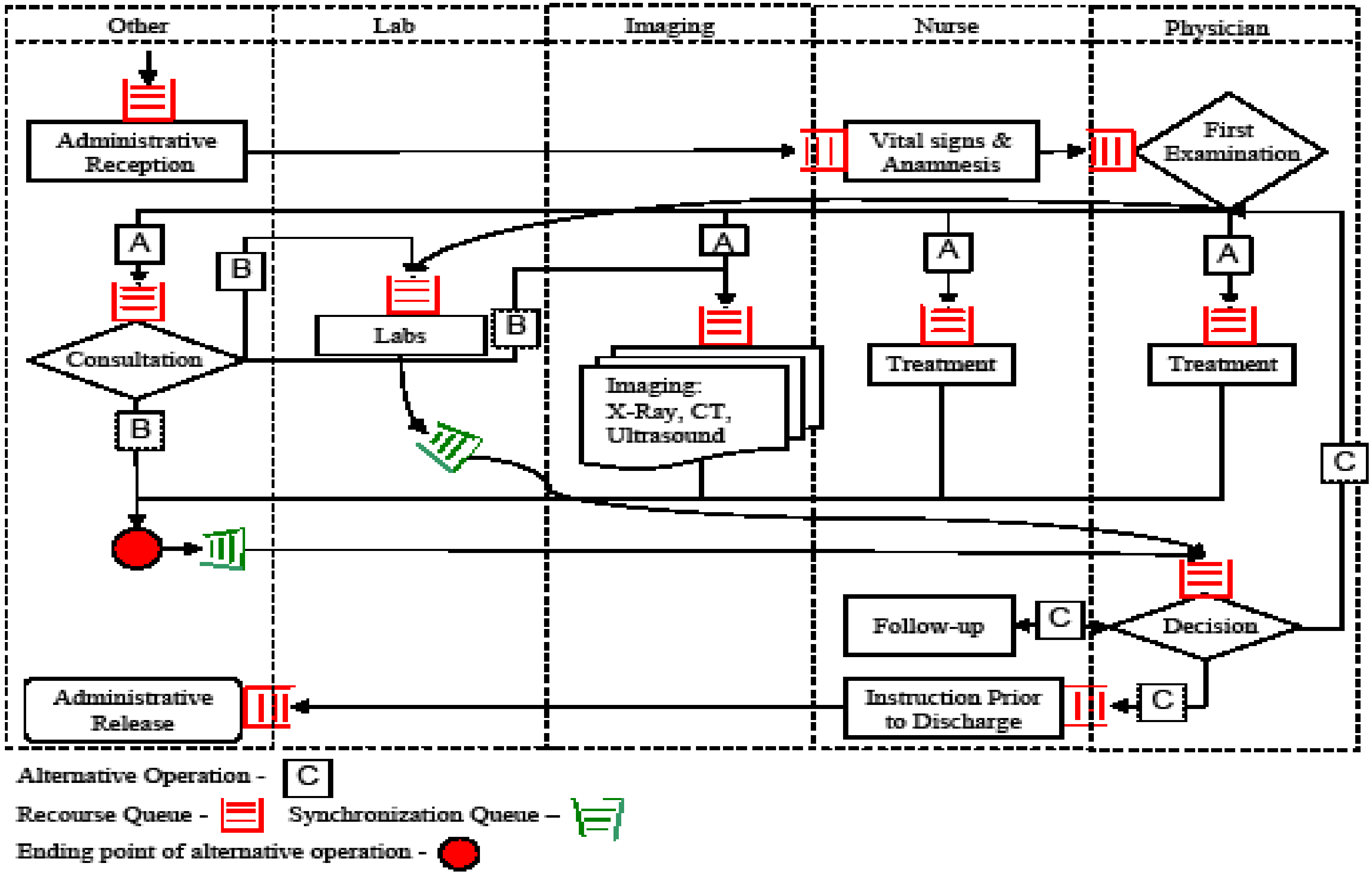
The Erlang A Model (*M*/*M*/*n*+*M*)
M for "Markov" – special assumptions



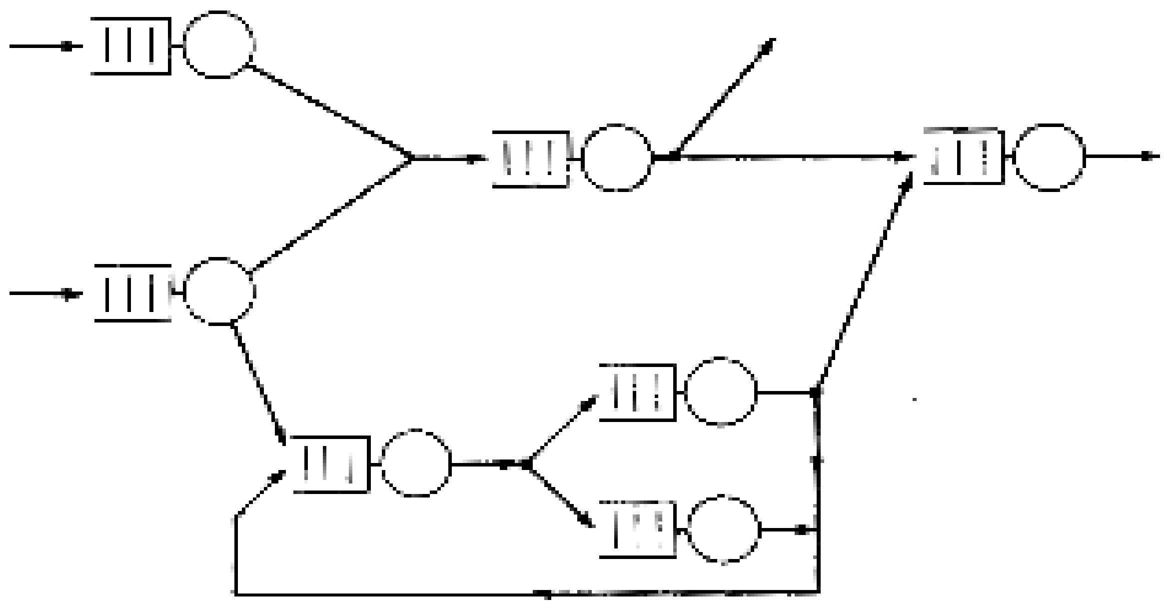
This week: Hospitals

Intensive Care Unit (ICU)

Patient Flow in an Emergency Department



View as
An Open Network of Queues



Jackson (Open) Markov Queueing Network

(Section 7 in CTMC notes, handout and Courseworks)

$X_j(t)$ = number of customers at station j at time t

$(X_1(t), \ldots, X_m(t))$ = **CTMC**

Assumptions:

- Customers come from outside and eventually leave
- Unlimited waiting room
- Number of stations is $\boldsymbol{m}$
- External Poisson arrivals at each queue, rate $\boldsymbol{\lambda_{e,j}}$
- Exponential service times at each station, rate $\boldsymbol{\mu_j}$
- $\boldsymbol{s_j}$ servers at station j
- Markovian routing, $\boldsymbol{(P_{i,j}) = P}$

Product-Form Steady-State Distribution

Steady state: **convergence in distribution**

$(X_1(t), \ldots, X_m(t)) \rightarrow (X_1(\infty), \ldots, X_m(\infty))$ as $t \rightarrow \infty$

Can find by **$\alpha Q = 0$,** but do not need to.

$P(X_1(\infty) = j_1, \ldots, X_m(\infty) = j_m) =$

$P(X_1(\infty) = j_1) \times \ldots \times P(X_m(\infty) = j_m)$

Product-form = independence

where station j is $M/M/s_j$ queue

Traffic Rate Equations

$$\lambda_j = \lambda_{e,j} + \sum_{i=1}^{m} \lambda_i P_{i,j} \quad \text{for} \quad 1 \le j \le m$$

$$\Lambda = \Lambda_e + \Lambda P$$ (matrix version)

$$\Lambda = \Lambda_e (I - P)^{-1}$$ (solution)

traffic intensity at station *j*: **$\rho_j = \lambda_j / s_j \mu_j$**

station *j* is $M/M/s_j$ queue

With parameters: **λ_j, μ_j** and **s_j**

The Queueing Network Analyzer (QNA)

(WW 1983, 1989 papers)

- ***Approximate*** Analysis of non-Markov OQN Model
- Characterize (quantify) ***variability*** at each queue
  - Arrival parameters: λ, c_a^2
  - Service parameters: μ, c_s^2
  - c^2 = Variance/Mean^2
- Characterize (quantify) impact on performance
- Parametric-Decomposition Method
  - Network calculus for variability parameters
- Approximate Performance of GI/GI/s queues

Variability Parameters

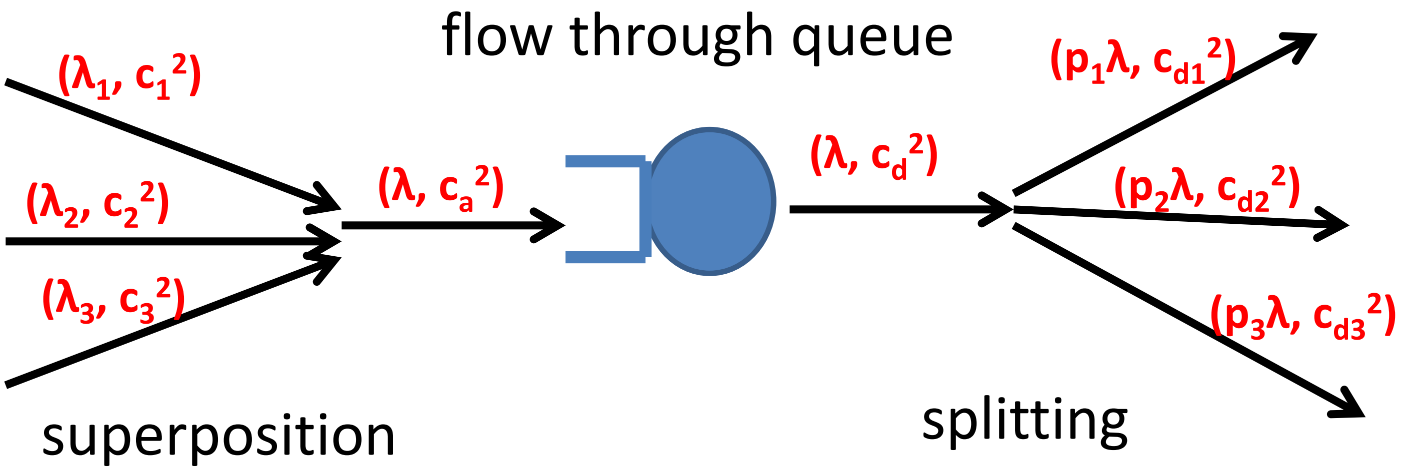
For service times and interarrival times, think of IID random variables

squared coefficient of variation

$$c^2 = c_X^2 = \mathrm{Var}(X)/E[X]^2$$

Exponential: $c^2 = 1$
Deterministic: $c^2 = 0$

Three Operations



From Routes to Markov Routing

- Start with routes for each class of customers
- Aggregate to convert to Markovian routing
  - P_{ij} = proportiion of departures from queue I that go next to queue *j*
  - ($1/\mu_j$) = average service time at queue *j*
  - c^2_{sj} = appropriate scv
    - Second moment average of second moments
    - Variance = second moment – mean^2
    - Scv = variance/ mean^2

Example 1, p. 2791 of 1983 paper

- *2* queues and *r = 3* routes

$$(n_k,\ \hat{\lambda}_k,\ c_k^2;\ n_{k1},\ \tau_{k1},\ c_{sk1}^2;\ \cdots;\ n_{kn_k},\ \tau_{kn_k},\ c_{skn_k}^2).$$

Here suppose that the *r* vectors are:

$$(2,\ 2,\ 1;\ 1,\ 1,\ 1;\ 1,\ 3,\ 3)$$

$$(3,\ 3,\ 2;\ 1,\ 2,\ 0;\ 2,\ 1,\ 1;\ 1,\ 2,\ 1)$$

$$(2,\ 2,\ 4;\ 2,\ 1,\ 1;\ 1,\ 2,\ 1).$$

A Manufacturing Example

(from p. 1151 of Segal and WW 1989)

Table 1. Comparisons Between QNA and Simulation

Model	Number of				WIP (Lots) Mean		Yield	Interval (Time Units) Mean		Interval (Time Units) Standard Deviation	
	Workstations	Products	Operations	Hops	QNA	SIM	QNA/SIM	QNA	SIM	QNA	SIM
1	67	1	135	9	261.6	255.5	1.0045	34.2	32.4	4.81	4.97
2	30	1	108	15	41.8	40.5	1.000	12.7	12.4	2.8	–

But could have fork-join: Last Class Arrest-to-Arraignment Process Larson et al. (1993)