Palm Calculus
Part 1
The Importance of the Viewpoint

JY Le Boudec

This fish is black

This fish is white
Who says the truth?

BorduRail: according to our systematic tracking system, proba of a train being late $\leq 5$

BorduKonsum: according to our consumer survey, proba of being late $\approx 30\%$
1. Event versus Time Averages

Consider a simulation, state \( S_t \)
Assume simulation has a stationary regime

Consider an *Event Clock*: times \( T_n \) at which some specific changes of state occur

- Ex: arrival of job; Ex. queue becomes empty

*Event average* statistic

\[
\bar{Q}^0 := \frac{1}{N + 1} \sum_{n=0}^{N} Q(T_n^-)
\]

*Time average* statistic

\[
\bar{Q} := \frac{1}{T_N - T_0} \int_{T_0}^{T_N} Q(s) \, ds
\]
Example: Gatekeeper; Average execution time

Two processes, with execution times 5000 and 1000

\[ W_s = \frac{5000 + 1000}{2} = 3000 \]

Inspector arrives at a random time

red processor is used with proba \( \frac{90}{100} \)

\[ W_c = \frac{90}{100} \times 5000 + \frac{10}{100} \times 1000 \]

\[ = 4600 \]
Sampling Bias

$W_s$ and $W_c$ are different

A metric definition should mention the sampling method (viewpoint)

Different sampling methods may provide different values: this is the sampling bias

*Palm Calculus* is a set of formulas for relating different viewpoints

Can often be obtained by means of the *Large Time Heuristic*
Large Time Heuristic Explained on an Example

We want to relate $W_s$ and $W_c$

We apply the large time heuristic

1. How do we evaluate these metrics in a simulation?

$$W_s = \frac{1}{N} \sum_{n=1}^{N} X_n = \bar{X}$$

$$W_c = \frac{1}{T} \int_{0}^{T} X_{N^+(t)} dt$$

where $N^+(t) =$ index of next green or red arrow at or after $T$
Large Time Heuristic Explained on an Example

2. Break one integral into pieces that match the $T_n$’s:

$$W_S = \frac{1}{N} \sum_{n=1}^{N} X_n = \bar{X}$$

$$W_c = \frac{1}{T} \int_0^T X_{N^+(t)} dt$$

$$W_c = \frac{1}{T} \left( \int_0^{T_1} X_{N^+(t)} dt + \int_{T_1}^{T_2} X_{N^+(t)} dt + \cdots + \int_{T_{N-1}}^{T_N} X_{N^+(t)} dt \right)$$

$$= \frac{1}{T} \left( \int_0^{T_1} X_1 dt + \int_{T_1}^{T_2} X_2 dt + \cdots + \int_{T_{N-1}}^{T_N} X_N dt \right)$$

$$= \frac{1}{T} \left( T_1 X_1 + (T_2 - T_1) X_2 + \cdots + (T_N - T_{N-1}) X_N \right)$$

$$= \frac{1}{T} \left( S_1 X_1 + S_2 X_2 + \cdots + S_N X_N \right)$$
Large Time Heuristic Explained on an Example

3. Compare

\[ W_c = \frac{1}{T} (S_1 X_1 + S_2 X_2 + \cdots + S_N X_N) \]

\[ = \frac{N}{T} \times \frac{1}{N} (S_1 X_1 + S_2 X_2 + \cdots + S_N X_N) \]

\[ = \lambda \times (\text{cov}(S, X) + \bar{S} \bar{X}) = \lambda \times \left( \text{cov}(S, X) + \frac{1}{\lambda} \bar{X} \right) \]

\[ W_c = \lambda \text{cov}(S, X) + W_s \]
This is Palm Calculus!

\[ W_c = \lambda \text{cov}(S, X) + W_s \]

**Viewpoint 1: System Designer**

Two processes, with execution times 5000 and 1000

\[ W_e = \frac{5000 + 1000}{2} = 3000 \]

**Viewpoint 2: Customer**

Inspector arrives at a random time red processor is used with proba \( \frac{90}{100} \)

\[ W_i = \frac{90}{100} \times 5000 + \frac{10}{100} \times 1000 = 4600 \]
In which case do we expect to see $W_c > W_s$?

A. $S_n = 90, 10, 90, 10, 90; X_n = 5000, 1000, 5000, 1000, 5000$
B. $S_n = 90, 10, 90, 10, 90; X_n = 1000, 5000, 1000, 5000, 1000$
C. Both
D. None
E. I don’t know
Solution

In case A, $S_n$ and $X_n$ are positively correlated (when the interval is long, so is the processing time), i.e. $\text{cov}(X, S) > 0$. By the Palm calculus formula: $W_c > W_s$

In case B, the correlation is negative, therefore $W_c < W_s$

Answer A
The Large Time Heuristic

1. formulate each performance metric as a long run ratio, as you would do if you would be evaluating the metric in a discrete event simulation;
2. take the formula for the time average viewpoint and break it down into pieces, where each piece corresponds to a time interval between two selected events;
3. compare the two formulations.

Formally correct if simulation is stationary

It is a robust method, i.e. independent of assumptions on distributions (and on independence)
Other «Clocks»

**Example 7.4: Flow versus Packet Clock [96].** Packets arriving at a router are classified in “flows”. We would like to plot the empirical distribution of flow sizes, counted in packets. We measure all traffic at the router for some extended period of time. Our metric of interest is the probability distribution of flow sizes. We can take a flow “clock”, or viewpoint, i.e. ask: pick an arbitrary flow, what is its size? Or we could take a packet viewpoint and ask: take an arbitrary packet, what is the magnitude of its flow? We thus have two possible metrics (Figure 7.3):

```
Distribution of flow sizes
for an arbitrary flow
for an arbitrary packet
```

```
Flow 1

Flow 2

Flow 3
```
Which curves are for the per-packet viewpoint?

A. A
B. B
C. It depends
D. I don’t know
Answer A
There are more packets in the large flows. So more packets experience a large flow size.
Load Sensitive Routing of Long-Lived IP Flows
Anees Shaikh, Jennifer Rexford and Kang G. Shin
Proceedings of Sigcomm'99

ECDF, per packet viewpoint

ECDF, per flow viewpoint
Per flow: $f_F(s) = \frac{1}{N} \times \text{number of flows with length } s$, where $N$ is the number of flows in the dataset;

Per packet: $f_P(s) = \frac{1}{P} \times \text{number of packets that belong to a flow of length } s$, where $P$ is the number of packets in the dataset;

Mean flow size:
per flow: $S_F$
per packet: $S_P$
Large «Time» Heuristic

1. How do we evaluate these metrics in a simulation?

per flow \( S_F = \frac{1}{N} \sum_n S_n \)

per packet \( S_P = \frac{1}{P} \sum_p S_{F(p)} \)

where \( F(p) = n \) when packet \( p \) belongs to flow \( n \)
Large «Time» Heuristic

1. How do we evaluate these metrics in a simulation?
   - per flow: $S_F = \frac{1}{N} \sum_n S_n$
   - per packet: $S_P = \frac{1}{P} \sum_p S_{F(p)}$

   where $F(p) = n$ when packet $p$ belongs to flow $n$

2. Put the packets side by side, sorted by flow

Flow n=1

Flow n=2

Flow n=3

Size $S_1$

Size $S_2$

Size $S_3$

$S_P = \frac{1}{P} \left( S_1 + S_1 + S_2 + S_2 + S_3 + S_3 + S_3 + S_3 + \cdots \right)$

$= \frac{1}{P} \left( S_1 \times S_1 + S_2 \times S_2 + S_3 \times S_3 + \cdots \right) = \frac{1}{P} \sum_n S_n^2$
Large «Time» Heuristic

3. Compare

\[ S_P = \frac{1}{P} \sum_n S_n^2 \]

\[ S_F = \frac{1}{N} \sum_n S_n = \frac{1}{N} P \]

\[ S_P = \frac{N}{P} \times \frac{1}{N} \sum_n S_n^2 = \frac{1}{S_F} \times \frac{1}{N} \sum_n S_n^2 = \frac{1}{S_F} \times \left( S_F^2 + \text{var}_F(S) \right) \]

\[ S_P = S_F + \frac{1}{S_F} \text{var}_F(S) \]
Large «Time» Heuristic for PDFs of flow sizes

Put the packets side by side, sorted by flow

1. How do we evaluate these metrics in a simulation?

1. For $s$ spanning the set of observed flow sizes:

$$f_F(s) = \frac{1}{N} \sum_{n=1}^{N} 1\{S_n = s\}$$ \hspace{1cm} (7.8)

$$f_P(s) = \frac{1}{P} \sum_{p=1}^{P} 1\{S_{F(p)} = s\}$$ \hspace{1cm} (7.9)

where $S_n$ be the size in bytes of flow $n$, for $n = 1, \ldots N$, and $F(p)$ is the index of the flow that packet number $p$ belongs to.
1. For $s$ spanning the set of observed flow sizes:

$$f_F(s) = \frac{1}{N} \sum_{n=1}^{N} \mathbf{1}_{\{S_n=s\}}$$  \hspace{1cm} (7.8)$$

$$f_P(s) = \frac{1}{P} \sum_{p=1}^{P} \mathbf{1}_{\{S_{F(p)}=s\}}$$  \hspace{1cm} (7.9)$$

where $S_n$ be the size in bytes of flow $n$, for $n = 1, \ldots N$, and $F(p)$ is the index of the flow that packet number $p$ belongs to.

2. We can break the sum in Eq.(7.9) into pieces that correspond to ticks of the flow clock:

$$f_P(s) = \frac{1}{P} \sum_{n=1}^{N} \sum_{p:F(p)=n} \mathbf{1}_{\{S_n=s\}} = \frac{1}{P} \sum_{n=1}^{N} \sum_{p=1}^{P} \mathbf{1}_{\{F(p)=n\}} \mathbf{1}_{\{S_n=s\}}$$

$$= \frac{1}{P} \sum_{n=1}^{N} \mathbf{1}_{\{S_n=s\}} \sum_{p=1}^{P} \mathbf{1}_{\{F(p)=n\}} = \frac{1}{P} \sum_{n=1}^{N} \mathbf{1}_{\{S_n=s\}} = \frac{s}{P} \sum_{n=1}^{N} \mathbf{1}_{\{S_n=s\}}$$  \hspace{1cm} (7.10)$$

3. Compare Eqs.(7.8) and (7.10) and obtain that for all flow size $s$:

$$f_P(s) = \eta s f_F(s)$$  \hspace{1cm} (7.11)$$

where $\eta$ is a normalizing constant ($\eta = N/P$).
(a) Empirical Complementary CDFs

(b) Histogram, flow viewpoint

(c) Histogram, packet viewpoint
Cyclist’s Paradox

On a round trip tour, there is more uphill than downhill.

Example 7.5: Kilometer versus Time Clock: Cyclist’s Paradox. A cyclist rides Swiss mountains; his speed is 10 km/h uphill and 50 km/h downhill. A journey is made of 50% uphill slopes and 50% downhill slopes. At the end of the journey, the cyclist is disappointed to read on his speedometer an average speed of only 16.7 km/h, as he was expecting an average of \( \frac{10 + 50}{2} = 30 \) \text{ km/h}. 
The km clock vs the standard clock

\( v_\ell = \text{speed for the } \ell^{\text{th}} \text{ kilometer} \)

\[
S_{\text{kilometer}} = \frac{1}{L} \sum_\ell v_\ell = \text{mean of } v_\ell
\]

\[
S_{\text{time}} = \frac{L}{T} = \frac{L}{\sum_\ell \frac{1}{v_\ell}} = \text{harmonic mean of } v_\ell < \text{mean of } v_\ell
\]

In this case the large time heuristic does not give a closed form relationship between the two averages; however, a closed form relationship can be obtained for the two distributions of speeds. Using the same method as in Example 7.4, one obtains

\[
f_t(v) = \eta \frac{1}{v} f_k(v) \tag{7.14}
\]

where \( f_t(v) \) [resp. \( f_k(v) \)] is the PDF of the speed, sampled with the standard clock [resp. km clock] and \( \eta \) is a normalizing constant; \( f_t \) puts more mass on the small values of the speed \( v \), this is another explanation to the cyclist’s paradox.
BorduRail claims that only 5% of trains arrivals are late

BorduKonsum claims that 30% of train users suffer from late train arrivals

A. At least one of them lies
B. The number of passengers in a late train $\approx 1.15 \times$ passengers in average train
C. The number of passengers in a late train $\approx 6 \times$ passengers in average train
D. I don’t know
Solution

\(N\) arrival events;
\(D_n = 1 \iff \text{arrival } n \text{ is late}, D_n = 0 \iff \text{arrival } n \text{ is on time}

\(P_n = \text{number of passengers leaving train at } n^{th} \text{ arrival}

BorduRail computes \(\bar{D} = \frac{1}{N} \sum_n D_n\)

BorduKonsum estimates \(D^* = \frac{1}{\sum_{n=1}^{N} P_n} \sum_{n=1}^{N} \sum_{p=1}^{P_n} D_n\)

Let \(\bar{P} = \frac{1}{N} \sum_{n=1}^{N} P_n \) (people per arrival)

\(N_{\text{late}} = \sum_{n=1}^{N} D_n = N \bar{D}\) and

\(\bar{P}_{\text{late}} = \frac{1}{N_{\text{late}}} \sum_n P_n D_n \) (people per late arrival)

So that: \(D^* = \frac{1}{\sum_{n=1}^{N} P_n} \sum_{n=1}^{N} P_n D_n = \frac{1}{N \bar{P}} N_{\text{late}} \bar{P}_{\text{late}} = \frac{1}{N \bar{P}} N \bar{D} \bar{P}_{\text{late}}\)

i.e. \(D^* = \frac{\bar{D}}{\bar{P}} \bar{P}_{\text{late}} \) Answer C.