



Balanced Sequences and Optimal Routing

Eitan Altman, Bruno Gaujal, Arie Hordijk

► **To cite this version:**

Eitan Altman, Bruno Gaujal, Arie Hordijk. Balanced Sequences and Optimal Routing. RR-3180, INRIA. 1997. inria-00073509

HAL Id: inria-00073509

<https://hal.inria.fr/inria-00073509>

Submitted on 24 May 2006

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Balanced Sequences and Optimal Routing

Eitan ALTMAN Bruno GAUJAL and Arie HORDIJK

N° 3180

Juin, 1997

———— THÈME 1 ————

 ***rapport
de recherche***

Balanced Sequences and Optimal Routing

Eitan ALTMAN^{*} Bruno GAUJAL^{**} and Arie HORDIJK^{***}

Thème 1 — Réseaux et systèmes
Projet Mistral, Sloop

Rapport de recherche n° 3180 — Juin, 1997 — 22 pages

Abstract: The objective pursued in this paper is two-fold. The first part gives an overview of the following combinatorial problem: is it possible to construct an infinite sequence over n letters where each letter is distributed as “evenly” as possible and appears with a fixed rate? The second objective of the paper is to relate this construction to the framework of optimal routing in queueing networks. We show under rather general assumptions that the optimal deterministic routing in several event graphs is such a sequence.

Key-words: Balanced sequences, multimodularity, event graphs, optimal control

(Résumé : tsvp)

^{*} INRIA, BP 93, 2004 Route des Lucioles, 06902 Sophia Antipolis Cedex, France. E-mail: altman@sophia.inria.fr. URL:<http://www.inria.fr:80/mistral/personnel/Eitan.Altman/me.html>

^{**} INRIA, BP 93, 2004 Route des Lucioles, 06902 Sophia Antipolis Cedex, France. E-mail: gaujal@sophia.inria.fr. Bruno Gaujal is a member of a common project between CNRS, UNSA and INRIA.

^{***} Dept. of Mathematics and Computer Science, Leiden University, P.O.Box 9512, 2300RA Leiden, The Netherlands. E-mail: hordijk@wi.leidenuniv.nl. The research of Arie Hordijk was done while he was on sabbatical leave at INRIA, Sophia-Antipolis; it has been partially supported by the Ministère Français de l'Éducation Nationale et de l'Enseignement Supérieur et de la Recherche.

Séquences équilibrées et routage optimal

Résumé : L'objectif de cet article est double. La première partie présente une vue d'ensemble sur le problème combinatoire suivant: est-il possible de construire une suite infinie sur n lettres dans laquelle chaque lettre est distribuée aussi régulièrement que possible et apparaît avec une fréquence fixée? Le deuxième objectif de cet article est de montrer l'intérêt d'une telle construction dans le cadre du routage optimal dans les réseaux de files d'attente. Nous montrons sous des conditions assez générales que le routage déterministe optimal est une suite équilibrée.

Mots-clé : suites équilibrées, multimodularité, graphes d'événements, contrôle optimal

1 Introduction

It is a rather general problem to consider a system with multiple resources and tasks. Tasks can be performed by any resource and arrive in the system sequentially. The problem is to construct a *routing* of the tasks to the resources to minimize a given cost function. Such models are common in multiprocessor systems and communication networks, where the cost function may be the combined load in the resources.

Here, we show that under rather general assumptions, the optimal routing sequence in terms of expected average workload in each resource is given by a *balanced sequence*, that is a sequence in which the option to route towards a given resource, is taken in an evenly distributed fashion.

This motivates the first part of the paper which is essentially based on word combinatorics and uses its specific vocabulary. This part relies heavily on results in [10] and also in [15, 16, 13].

The second part will show an application of balanced sequences for routing problems. This part uses results from convex analysis, mainly from [9, 2, 1]. It shows the link between balanced sequences and scheduling problems, such as routing among several systems.

The paper is structured as follows. In the second section, we introduce a formal definition of *balanced sequence* and we present an overview on their properties. The section 3 makes the link between the notion of balanced sequences and the optimal scheduling in networks and is also devoted to prove the optimality of balanced words for routing customers in a multiple queue system. Section 4 presents special cases for which the optimal rates can be computed.

2 Balanced sequences

In this section, we will present the notion of balanced sequences, which is closely related to the notion of *Sturmian* sequences [13]. This presentation is not exhaustive and many other related articles can be consulted for further investigation on this topic [17, 10, 3, 5, 15]. Although the section is self contained and presents several result of interest by their own, we mostly focus the *rate problem* (see Problem 2), which will be used in the application section (§ 3).

2.1 Preliminaries

Let \mathcal{A} be a finite alphabet and $\mathcal{A}^{\mathbb{N}}$ the set of sequences defined on \mathcal{A} .

If $u \in \mathcal{A}^{\mathbb{N}}$, then a *word* W of u is a finite subsequence of consecutive letters in u : $W = u_a u_{a+1} \cdots u_{a+k-1}$. The integer k is the *length* of W and will be denoted $|W|$.

If $a \in \mathcal{A}$, $|W|_a$ is the number of a 's in the word W .

Definition 2.1. *The sequence $u \in \mathcal{A}^{\mathbb{N}}$ is balanced if for any two words W and W' in u of same length, and any $a \in \mathcal{A}$,*

$$-1 \leq |W|_a - |W'|_a \leq 1.$$

If $a \in \mathcal{A}$, we also define the *support* in u of the letter a as the function $\delta_a(u) : \mathcal{A}^{\mathbb{N}} \rightarrow \{0, 1\}^{\mathbb{N}}$ by, $\delta_a(u)_i = 1$ if $u(i) = a$ and 0 otherwise.

Definition 2.2. A sequence $d \in \{0, 1\}^{\mathbb{N}}$ is regular if there exists two real numbers θ and p such that $d_n = \lfloor (n+1)p + \theta \rfloor - \lfloor np + \theta \rfloor$.

We can establish the following relation between regular and balanced words.

Theorem 2.3. A sequence $u \in \mathcal{A}^{\mathbb{N}}$ is balanced if and only if $\delta_a(u)$ is regular for all $a \in \mathcal{A}$.

Proof. First, let us prove that $\delta_a(u)$ regular implies $-1 \leq |W|_a - |W'|_a \leq 1$ for any words W and W' in u . The interval $]k, k+m]$ in u contains exactly $e = \lfloor \theta + pk + pm \rfloor - \lfloor \theta + pk \rfloor$ times the letter a . Now, $\lfloor pm \rfloor + \lfloor \theta + pk \rfloor - \lfloor \theta + pk \rfloor \leq e \leq \lceil pm \rceil + \lfloor \theta + pk \rfloor - \lfloor \theta + pk \rfloor$. This shows the value of e can differ by at most one when k varies so u is balanced in a , and this holds for all a .

Conversely, it is shown in [7] using combinatoric arguments that if for a given letter a , $-1 \leq |W|_a - |W'|_a \leq 1$ then the sequence $\delta_a(u)$ is characterized by two real numbers, θ and p and we have: $\delta_a(u)_n = 0$ if $\theta + np \bmod 1 \in [0, 1-p)$ and 1 otherwise.

At this point, an easy induction shows that $\delta_a(u)_n = \lfloor (n+1)p + \theta \rfloor - \lfloor np + \theta \rfloor$. □

In [12], the sequence u is said to have the reduction property if $\delta_u(a)$ is regular for all $a \in \mathcal{A}$. Theorem 2.3 shows that u has the reduction property if and only if it is balanced.

2.2 Constant gap sequences

Constant gap sequences are strongly balanced sequences, in the following sense.

Definition 2.4. A sequence G is constant gap if for any letter a , $\delta_a(G)$ is periodic, with a period of the form $0 \cdots 010 \cdots 0$.

Note that this explains the fact that G is said to have constant gaps for the letter a , since each a is separated from the next a in G by a constant number of letters.

Proposition 2.5. Constant gap sequences are balanced.

Proof. For each letter a , $\delta_a(G)$ is of the form $(0^n 10^m)^\omega$. Therefore, $\delta_a(G)$ is regular with $p = 1/(m+n+1)$ and $\theta = m/(m+n+1)$. Using the characterization of balanced sequences given in Theorem 2.3, this shows that G is balanced. □

Proposition 2.6. Constant gap sequences are periodic.

Proof. For each letter a , $\delta_a(G)$ is periodic of period p_a . The period of G is $\text{lcm}(p_a, a \in \mathcal{A})$. □

In the next lemma, we give a characterization of constant gap sequences that stresses the fact that constant gap is some kind of strong balance.

Proposition 2.7. *G is constant gap if and only if, for any two finite words, W and W' included in G with $||W| - |W'| \leq 1$, then for each letter a , $||W|_a - |W'|_a \leq 1$.*

Proof. Let a be a letter in the alphabet.

First, assume that G is constant gap. Let us call l_a the gap between two a in G . If $||W|_a - |W'|_a \geq 2$, then, necessarily, $||W| - |W'| \geq 2$. Conversely, let $W = aUa$ and $W' = aU'a$ be any two words in G with no a in the subwords U and U' . If $||U| \geq |U'| + 1$, then we have $||U| - |U'| \leq 1$ and $||W|_a - |W'|_a = 2$. This is a contradiction. Therefore $||U| = |U'|$ and G is constant gap. \square

Since a constant gap sequence is balanced, each letter appears with a given rate in the sequence. note however that since a constant gap word is necessarily periodic, the rate of each letter is rational.

As we will do in section 2.4 for the case of balanced words, we now address the following question:

Problem 1: Given a set (q_1, \dots, q_N) , with $q_1 + \dots + q_N = 1$ is it possible to construct a constant gap word on N letters with rates (q_1, \dots, q_N) ?

We will not solve this problem for a general N . We will only give some properties of the set (q_1, \dots, q_N) which will be useful in the following. A characterization of such (q_1, \dots, q_N) is given in [18].

As a general remark, note that (q_1, \dots, q_N) are rational numbers of the form $q_i = k_i/d$, with d the smallest period of G . Therefore, we have, $\sum_i k_i/d = 1$ and for each i , k_i divides d . By definition of the rates, we also have $q_a = 1/p_a$ for all letters.

This problem has been addressed in [18] under a different point of view where the name of exact covering sequences was used instead of balanced sequences.

Definition 2.8. *The couples $\{(\theta_i, g_i), i = 1 \dots N\}$ is called an exact covering sequence if for every nonnegative integer n , there exists one and only one $1 \leq i \leq N$ such that $n = \theta_i \pmod{g_i}$.*

We have the following result.

Proposition 2.9. *The rates (q_1, \dots, q_N) are constant gap if there exists N numbers called phases, $\theta_1, \dots, \theta_N$ such that the couples $\{(\theta_i, p_i), i = 1 \dots N\}$ form an exact covering sequence.*

Proof. This property is a simple rewriting of the fact that each letter a_i in a constant gap sequence appears every $\{\theta_i + kp_i, k \in \mathbb{N}\}$. \square

Now, suppose that $\{(\theta_i, p_i), i = 1 \dots N\}$ is an exact covering sequence. Then in the series

$$S(x) \stackrel{\text{def}}{=} \sum_{i=1}^k \sum_{k \geq 0} x^{\theta_i + kp_i},$$

the coefficient of x^n in this series is equal to 1, $\forall n \geq 0$. Therefore, we have

$$S(x) = \sum_{i=1}^k \frac{x_i^\theta}{1 - x^{p_i}} = \frac{1}{1 - x}.$$

Using this characterization we have the following interesting property.

Lemma 2.10. *Assume $\{(\theta_i, p_i), i = 1 \dots N\}$ is an exact covering sequence and that $P = \max_i p_i$. Then P appears at least twice in the set p_1, \dots, p_N .*

Proof. The proof given here is similar to the discussion in [18] on the exact covering sequences. Let $w \stackrel{\text{def}}{=} e^{2i\pi/r}$ for some integer $r > 1$. By definition, w is a primitive r -th root of one. We have:

$$(w - x)S(x) = \sum_{i=1}^k \frac{(w - x)x_i^\theta}{1 - x^{p_i}} = \frac{w - x}{1 - x}.$$

Let $x \rightarrow w$. This yields

$$\sum_{i:r|p_i} \frac{-w^{\theta_i}}{-p_i w^{p_i-1}} = 0. \quad (1)$$

Now, take $r = P$. The set $\{i : P|p_i\}$ is exactly the set $\{i : P = p_i\}$. Equation (1) specified for $r = P$ can be written

$$\sum_{i:p_i=P} w^{\theta_i} = 0.$$

This implies that the set $\{i : P = p_i\}$ cannot be reduced to a single point since w is not zero. \square

To give some concrete examples, we consider the cases where N is small. First, note that in the case where the k_i are not all equal, (assume k_1 is the largest of all), we have

$$\sum_{i \neq 1} k_i/k_1 = d/k_1 - 1 = l_1,$$

where l_1 is the gap between two letters a_1 . This implies,

$$l_1 \leq N - 2. \quad (2)$$

Proposition 2.11. *There exists a constant gap sequence G with rates $(p, 1-p)$ if and only if $p = 1/2$.*

Proof. Let a be a letter in G with gap l . Since the alphabet contains only two letters, $l = 1$. This means $p = 1/2$. \square

Proposition 2.12. *There exists a constant gap sequence G with rates (p_1, p_2, p_3) if and only if the following holds: $(p_1, p_2, p_3) = (1/3, 1/3, 1/3)$ or $(1/2, 1/4, 1/4)$ (up to a permutation).*

Proof. Assume that $(p_1, p_2, p_3) \neq (1/3, 1/3, 1/3)$ (otherwise, $G = (abc)^\omega$ is constant gap). Using Equation (2), $l_1 = 1$ and $p_1 = 1/2$. Therefore, the sequence obtained from G when removing all the letters a_1 is constant gap. Applying Lemma 2.11 shows that $p_2 = p_3$. The only solution is $p_1 = 1/2, p_2 = 1/4$ and $p_3 = 1/4$. The associated constant gap sequence is $(a_1 a_2 a_1 a_3)^\omega$. \square

Proposition 2.13. *There exists a constant gap sequence G with rates (p_1, p_2, p_3, p_4) if and only if $\{p_1, p_2, p_3, p_4\}$ belong to the set (up to a permutation),*

$$\{(1/4, 1/4, 1/4, 1/4), (1/2, 1/4, 1/8, 1/8), (1/2, 1/6, 1/6, 1/6), (1/3, 1/3, 1/6, 1/6)\}.$$

Proof. If the rates are all equal, then $(p_1, p_2, p_3, p_4) = (1/4, 1/4, 1/4, 1/4)$. Now, note that Equation 2 implies that if the rates are not all equal $l_1 \leq 2$. We consider any letter $a_i, i \neq 1$, assume that the number of a_1 in between two a_i is not constant and takes values m and n , $m > n$. then we have $l_i \geq (n-1)l_1 + n$ on one hand and $l_i \leq n(l_1) - 3$ on the other hand. This is impossible since $l_1 \leq 2$. Therefore, the number of a_1 in between two a_i is constant. This is true for all i . The sequence formed by removing all a_1 is still constant gap. It has rates of the form $(1/3, 1/3, 1/3)$ or $(1/2, 1/4, 1/4)$. From this point a case analysis shows that the original word has rates $(1/2, 1/4, 1/8, 1/8), (1/2, 1/6, 1/6, 1/6)$ or $(1/3, 1/3, 1/6, 1/6)$ by inserting the letter a_1 in a constant gap sequence over the letters a_2, a_3, a_4 . \square

These few examples of constant gap sequences illustrate the fact that the rates in these sequences have very strong constraints.

2.3 Characterization of balanced sequences

Several studies have been recently done on balanced sequences [12, 15, 16, 8]. In [10], a characterization involving constant gap sequences is given.

Proposition 2.14 (Hubert). *Let U be a balanced sequence on the alphabet $\{0, 1\}$. Construct a new sequence S by replacing in U , the subsequence of zeros by a constant gap sequence G on an alphabet \mathcal{A}_1 , and the subsequence of ones by a constant gap sequence H on a disjoint alphabet \mathcal{A}_2 . Then S is balanced on the alphabet $\mathcal{A}_1 \cup \mathcal{A}_2$.*

Proof. We give a proof similar to Hubert's proof. Let a be a letter in \mathcal{A}_1 (the proof is similar for a letter in \mathcal{A}_2). Let W and W' be two words of S of the same length. Then, the corresponding words X and X' in U verify $||W|_0 - |W'|_0| \leq 1$ since U is balanced. If we keep only the 0's in X and X' , then the corresponding Z and Z' words in G satisfy $||Z| - |Z'|| \leq 1$. Since G is constant gap, and using Lemma 2.7, $||Z|_a - |Z'|_a| \leq 1$. We end the proof noting that the construction of Z and Z' implies $|Z|_a = |W|_a$ and $|Z'|_a = |W'|_a$. \square

Conversely, we have the following theorem.

Theorem 2.15 (Hubert). *Let $u \in \mathcal{A}^{\mathbb{N}}$ be balanced and non periodic. Then there exists a partition of \mathcal{A} into two sets \mathcal{A}_1 and \mathcal{A}_2 such that the sequence v defined by:*

$$v_n = 1 \quad \text{if } u_n \in \mathcal{A}_1, \quad (3)$$

$$v_n = 0 \quad \text{if } u_n \in \mathcal{A}_2, \quad (4)$$

is regular. Furthermore, the sequences z_1 and z_2 constructed from u by keeping only the letters from \mathcal{A}_1 and \mathcal{A}_2 respectively have constant gaps.

The proof of this theorem is technical and can be found in [10]. It is not reported here.

2.4 Rates in balanced sequences

Now, also note that a balanced sequence has several asymptotic properties, such as the following lemma.

Proposition 2.16. *A balanced sequence S has an asymptotic rate for any letter $a \in \mathcal{A}$, that is:*

$$\lim_{n \rightarrow \infty} \frac{|S_{[1,n]}|_a}{n} = p.$$

Proof. Since $|S_{[1,n]}|_a \leq n$, then $\frac{|S_{[1,n]}|_a + m}{n+m} \leq \frac{|S_{[1,n]}|_a}{n}$. Now,

$$\begin{aligned} \left| \frac{|S_{[1,n]}|_a}{n} - \frac{|S_{[1,n+m]}|_a}{n+m} \right| &\leq \frac{|S_{[n+1,n+m]}|_a + m}{n} \\ &\leq \frac{|S_{[1,m-1]}|_a + m + 1}{n}. \end{aligned}$$

When n goes to infinity, this quantity goes to zero, for all m . This proves that there exists p such that $\lim_{n \rightarrow \infty} \frac{\#(S_{[1,n]})}{n} = p$. □

Note that we have given an elementary proof of the existence of the asymptotic rate. Using Theorem 2.3, it is easy to see immediately that the asymptotic rate does exist and is equal to the real number p given in the definition of regular sequences.

Next, we formulate the problem which we will study in this section:

Problem 2: Given a set (p_1, \dots, p_N) , is it possible to construct a balanced word on N letters with rates (p_1, \dots, p_N) ?

We will see in the following that this construction is not possible for all the values of the rates (p_1, \dots, p_N) . If a n -tuple (p_1, \dots, p_N) makes the construction possible, the tuple is said to be *balanced* as well.

2.5 The case $N = 2$

This case is well known and balanced words with two letters have been extensively studied (see for example [6, 13]). The following result is known even if it is often given under different forms.

Theorem 2.17. *For all p , $0 \leq p \leq 1$, the set of rates $(p, 1 - p)$ is balanced.*

Proof. The proof is similar to the proof of the first part of Theorem 2.3. We construct a sequence S as the support of the function $s(n) = \lfloor pn \rfloor - \lfloor p(n-1) \rfloor$. S is a balanced sequence because the interval $]k, k+m]$ contains exactly $e = \lfloor pk + pm \rfloor - \lfloor pk \rfloor$ elements of S . Now, $\lfloor pm \rfloor + \lfloor pk \rfloor - \lfloor pk \rfloor \leq e \leq \lceil pm \rceil + \lfloor pk \rfloor - \lfloor pk \rfloor$. This shows the value of e can differ by at most one when k varies so S is a balanced sequence. If S' is the complementary set of S , then it should be clear that S' has asymptotic rate $1 - p$ and S' is balanced because $S'_{]k, k+m]}$ contains $m - e$ elements. \square

Note that the proof of Theorem 2.17 also gives a construction of a balanced word with the given rates.

2.6 The case $N = 3$

The case $N = 3$ is essentially different from the case $N = 2$. In the case $N = 2$, all possible rates are balanced while when $N = 3$, there is essentially only one set of rates which is balanced. This result was partly proved and conjectured in [12] and consequently proved in [15].

Theorem 2.18 (Tijdeman). *A set of rates (p_1, p_2, p_3) is balanced if and only if,*

$$(p_1, p_2, p_3) = (4/7, 2/7, 1/7)$$

or two rates are equal.

Proof. We give a proof slightly simpler than the proof in [15]. First, assume that $p_1 = p_2$. Then, let S be a balanced sequence with two letters $\{a, b\}$ constructed with the rates $(p_1 + p_1, p_3)$. In S , replace alternatively the “a”s by the letters a_1, a_2 , we get a word S' on alphabet $\{a_1, a_2, b\}$ with rates (p_1, p_1, p_3) . Let us show that S' is balanced. Since S is balanced, the number of “a”s in an interval of length m is k or $k + 1$, for some k . Now, for S' , the number of “ a_1 ”s (resp. “ a_2 ”s) in such an interval is either $(k - 1)/2$ or $(k + 1)/2$ if k is odd and $k/2$ or $k/2 + 1$ if k is even. This proves that S' is balanced.

Now, assume that (p_1, p_2, p_3) are three different number. We assume that $p_1 > p_2 > p_3$. We will try to construct a sequence W with these respective rates on the alphabet $\{a, b, c\}$.

step 1: the sequence “aca” must appear in W .

There exists a pair of consecutive “a” with no “b” in between since $p_1 > p_2$. This means that the sequences “aa” or “aca” appear. If “aa” appears, then a “c” is necessarily surrounded by two “a”s.

step 2: the sequence “baab” must appear in W .

There exist a pair of consecutive “b” with no “c” in between. This sequence is of the form “ $ba^n b$ ”. Now, $n \leq 1$ is not possible because of the presence of “aca” and b -regularity. $n \geq 3$ implies the existence of “ $a^{n-1}ca^{n-1}$ ” by a -regularity which is incompatible with “ $ba^n b$ ” because of b -regularity. Therefore $n = 2$. Note that this also implies the existence of “aa” and of “abaaba”.

step 3: the sequence “abacaba” appears in W .

the word W must contain a “c”. This “c” is necessarily surrounded by two “a”s since “aa” exists by a -regularity. This group is necessarily surrounded by two “b”s since “baab” exists, necessarily surrounded by two “a”s. We get the sequence “abacaba”.

Last step: $W = (abacaba)^\omega$.

No letter around this sequence can be a “c” because a sequence with four letters and no “c” exists: “baab”. None can be a “b” since a sequence with three letters and no “b”, namely “aca” exists. Therefore, they have to be two “a”s. Then note that the following two letters cannot be a “c” (because of the existence of “abaaba”), nor an “a” (because of the existence of “bac”) so it is a “b”, then followed necessarily by an “a”. At this point, we have the sequence “ $\star abacabaaba \star$ ”. Both \star are necessarily “c”s.

To end the proof, note that a “c” necessarily has the sequence “abaaba” on its right and the same sequence on its left. Finally note that the sequence “abaaba” is necessarily surrounded by two “c”s. Therefore the word W is periodic of the form $(abacaba)^\omega$. \square

2.7 The case $N = 4$

The case $N = 4$ is very similar to the case $N = 3$ for distinct rates. However when two rates are equal, this case is more complicated.

Theorem 2.19. *A rate tuple (p_1, p_2, p_3, p_4) with four distinct rates is balanced if and only if $(p_1, p_2, p_3, p_4) = (8/15, 4/15, 2/15, 1/15)$.*

Proof. We suppose that $p_1 > p_2 > p_3 > p_4$ and we show that there is only one balanced word with frequencies $p_1 > p_2 > p_3 > p_4$ and those frequencies are $(8/15, 4/15, 2/15, 1/15)$.

The proof involves different steps.

Step 1: W contains the sequences “aca” or “ada” or “acda” or “adca”.

There exists two consecutive “a”s with no “b” in between because $p_1 > p_2$. Therefore, either “aa” or “aca” or “ada” or “acda” or “adca” exist. If “aa” exists, then, a “c” is surrounded by two “a”s.

Step 2: W contains the sequence “baab”

First, we show that if a sequence “ $ba^n b$ ” exists, then $n = 2$. Indeed, the fact that “ada” or “aca” or “acda” or “adca” exist makes “bb” and “bab” impossible. On the other hand, if $n \geq 3$, the existence of “ $a^{n-1}ca^{n-1}$ ” is necessary by a -regularity and is incompatible with the existence of “ $ba^n b$ ” because of b -regularity.

Now, if no sequence of the form “ $ba^n b$ ” exists, then there exist two consecutive “b”s with one “d” and no “c” in between. This sequence (s_1) is of the form: “ $a^i b a^j d a^k b a^l$ ”. Note that the numbers i, j, k, l may be equal to zero but $i + j + k + l \geq 1$ by b -regularity.

There also exist two consecutive “ c ”s with no “ d ” in between. This sequence (s_2), is of the form: “ $ca^m ba^n c$ ”. Note that i, j, k, l, m, n are integers that can differ by at most one, the length of s_1 , $\text{len}(s_1) \geq \max(4, 2(n+m)+1) \geq n+m+3 = \text{len}(s_2)$. This is impossible by c -regularity.

step 3: the sequence “abacabaabacaba” exists in W . There exists two consecutive “ c ”s with no “ d ” in between. from the proof of Lemma 2.18, we know that a “ c ” is necessarily surrounded by the sequence “ aba ”. Therefore we have: “abacabaabaUcaba”, where U is a sequence that contains no “ d ” and no “ c ”. U cannot start with an “ a ” (because of “bacab”) cannot start with a “ b ” (because of “aca”) cannot start with a “ c ” and cannot start with a “ d ” by construction. Therefore U has to be empty.

Step 4: W is uniquely defined and is periodic of period “abacabadabacaba”: somewhere, W contains a “ d ”. From this point on, we can extend the sequence uniquely as: “abacabadabacaba” around this “ d ”, and the sequence “abacabaabacaba” has to be surrounded by two “ d ”s. This ends the proof. \square

To complete the picture, it is not difficult to see that

Proposition 2.20. *If the tuple (p_1, p_2, p_3, p_4) is made of less than two distinct numbers, then it is balanced.*

Proof. First, if the rates are all equal, they are obviously balanced. If three of them are equal, say $p_1 = p_2 = p_3$, then, we can construct a balanced word with rates $(3p_1, p_4)$ and we construct a balanced word with rates (p_1, p_1, p_1, p_4) by using Proposition 2.14 (where we take $G = \{p_1, p_2, p_3\}$, and $H = \{p_4\}$). If two pairs of rates are equal, say $p_1 = p_2$ and $p_3 = p_4$, then we construct a balanced word with rates $(2p_1, 2p_3)$ and we apply Proposition 2.14 twice, once for each sequence. \square

If the tuple (p_1, p_2, p_3, p_4) is made of exactly three distinct numbers, then this is a more complex case which is not studied here.

2.8 The general case

In this section, we are interested in the case of arbitrary N . First, note that Proposition 2.20 easily generalizes to any dimension.

Proposition 2.21. *If the tuple $(p_1, p_2, p_3, \dots, p_N)$ is made of less than two distinct numbers, then it is balanced.*

Proof. The proof is similar to the proof of Proposition 2.20. \square

Proposition 2.22. *If (p_1, p_2, \dots, p_N) is balanced, then $(\underbrace{p_1/k, \dots, p_1/k}_k, p_2, \dots, p_N)$ is balanced.*

Proof. The proof is very similar to the first part of Theorem 2.18. If W is a balanced word with letters $\{a_1, \dots, a_N\}$, consider the word W' constructed starting from W and replacing alternatively each “ a_1 ” by “ b_1 ”, “ b_2 ”, \dots , “ b_k ”. Note also that W' has the following set of rates, $(p_1/k, \dots, p_1/k, p_2, \dots, p_N)$.

Next, we show that W' is balanced. Since W is balanced, for an arbitrary integer m , the number of “ a_1 ”s in an interval of length m is n or $n+1$. Now, for W' , the number of “ b_i ”s in such an interval is either $\lfloor (n-1)/k \rfloor$ or $\lfloor (n+1)/k \rfloor$. This proves that W' is balanced. \square

For the general case and distinct rates, it is natural to give the following conjecture:

Conjecture 2.23. *A set of distinct rates $\{p_1, \dots, p_N\}$ is balanced if and only if*

$$\{p_1, \dots, p_N\} = \{2^{N-1}/(2^N - 1), \dots, 2^{N-i}/(2^N - 1), \dots, 1/(2^N - 1)\}.$$

We have not been able to prove this fact. We only have partial results given in the following lemmas.

Lemma 2.24. *The sequence $(2^{N-1}/(2^N - 1), \dots, 2^{N-i}/(2^N - 1), \dots, 1/(2^N - 1))$ is balanced, for all $N \in \mathbb{N}$.*

This lemma is the “if” direction of the conjecture.

Proof. We construct a balanced word W_N in the following inductive way: We assume that W_{N-1} is periodic of period $2^{N-1} - 1$, $W_{N-1} = V^\omega$ where V is a finite word in the alphabet $\{a_1, \dots, a_{N-1}\}$ and we construct $W_N = (Va_NV)^\omega$. First note that W_N has rates $(2^{N-1}/(2^N - 1), \dots, 2^{N-i}/(2^N - 1), \dots, 1/(2^N - 1))$. Then, we show that W_N is balanced by induction. For the word W_N , the support for all letters (say letter $N - i$) is periodic with the same period and is of the form of $2^N - j - 1$ intervals of the same length (2^j) except one of length $2^j - 1$.

By construction of W_N , this properties still hold and therefore, W_N is balanced. \square

Lemma 2.25. *Let W be balanced with rates $p_1 > \dots > p_N$, then, W is periodic. In particular, this means that $p_i \in \mathbb{Q}$, $\forall 1 \leq i \leq N$.*

Proof. If W is not periodic, Theorem 2.15 says that W is composed of two constant gap words. At least one of these words has at least two letters, and therefore two letters have rates which are equal by Lemma 2.10. Therefore, the rates in W of these two letters are also equal. \square

Lemma 2.26. *Let W be balanced with rates $p_1 > \dots > p_N$, with the following property: for any $1 \leq i \leq N$, there exists two consecutive letters “ a_i ” with no a_j in between, with $j > i$. Then, $(p_1, \dots, p_N) = (2^{N-1}/(2^N - 1), \dots, 2^{N-i}/(2^N - 1), \dots, 1/(2^N - 1))$.*

This lemma is a partial “only if” result for the conjecture.

Proof. The proof holds by induction. V_k denotes the period of the balanced word with rates $(2^{k-1}/(2^k - 1), \dots, 2^{k-i}/(2^k - 1), \dots, 1/(2^k - 1))$. We recall that according to the proof of lemma 2.24, $V_k = V_{k-1}a_kV_{k-1}$. We will prove by induction that W is periodic of period V_N .

We first prove that W contains the sequence $V_{k-1}a_kV_{k-1}V_{k-1}a_kV_{k-1}$ and that for the sequence V_{k-1} and that each letter in W , “ a_j ”, $j \geq k$ is surrounded by V_{k-1} , for all possible k . Now, for the first step of the induction ($k = 1$), note that according to the property on W , W contains the sequence “ a_1a_1 ” which is the same as “ V_1V_1 ”. Therefore any other letter is surrounded by two “ a_1 ”s. This also implies the existence of “ $a_1a_2a_1a_1a_2a_1$ ” which ends the case $k = 1$.

For the general case, by the induction assumption, a_k is surrounded by V_{k-2} , and we have the sequence: “ $V_{k-2}a_kV_{k-2}$ ” The existence of “ $V_{k-2}a_{k-1}V_{k-2}V_{k-2}a_{k-1}V_{k-2}$ ” proves that this sequence is surrounded with two “ a_{k-1} ”s. Therefore, two consecutive a_k form the sequence:

$$V_{k-2}a_{k-1}V_{k-2}a_kV_{k-2}a_{k-1}V_{k-2}UV_{k-2}a_{k-1}V_{k-2}a_kV_{k-2}a_{k-1}V_{k-2},$$

where U does not contain any letter “ a_j ”, $j \geq k$. If U contains “ a_{k-1} ”, then U is “ $V_{k-2}a_{k-1}V_{k-2}$ ”. This is impossible because of the existence of “ $a_{k-1}V_{k-2}a_kV_{k-2}a_{k-1}$ ”. If U contains any other letter “ a_i ”, then U is reduced to this letter, and by construction of V_{k-2} , the presence of “ $a_iV_{i-1}a_i$ ” contradicts the existence of “ $a_iV_{i-1}V_{i-1}a_i$ ”. Therefore, U is empty and we have the second part of the proof.

Now, we finish the proof by noticing that the letter a_N is surrounded by V_{N-1} and by noting that V_{N-1} is necessarily surrounded by “ a_N ”. \square

Lemma 2.27. *Assume that W is balanced. Assume that $p_a \geq 0.5$. Then the projection W' of W over the alphabet $\mathcal{A} - \{a\}$ is also balanced.*

Proof. Choose two sequences V'_1, V'_2 of length n in W' . Let V_1 and V_2 be any two sequences in W whose projections over the alphabet $\mathcal{A} - \{a\}$ are V'_1 and V'_2 , respectively. Assume, furthermore, that the first and last letters in V_1 and V_2 are not a . Let $k = |V_1| - n$ and $l = |V_2| - n$ denote the number of appearances of the letter a in V_1 and V_2 , resp.

Step 1: If $l = k$ then the difference in the number of occurrences of any letter b in V'_1 and in V'_2 is at most 1, since W is most regular, and since the number of b 's in V_1 (resp. V_2) is the same as its number in V'_1 (resp. V'_2).

Step 2: Assume that $l > k + 1$.

- Let \hat{V}_2 be the word obtained from V_2 by truncating the first and last letter. Then $|\hat{V}_2| = n + l - 2$, and the number of a 's in \hat{V}_2 is l .

- Let \hat{V}_1 be the word obtained from V_1 by adding to it the next $m = l - k - 2$ letters that appears after V_1 in the sequence W . Then $|\hat{V}_1| = n + l - 2 = |\hat{V}_2|$, and the number of a 's in \hat{V}_1 is not larger than $k + m = l - 2$. This is a contradiction with the fact that W is balanced.

Step 3: It remains to check $l = k + 1$. Add to V_2 the next letter that occurs in W to its right, to form the new word \bar{V}_2 . If it is not a then we have two successive letters that are

not a , which contradicts the fact that a has an asymptotic frequency of at least $1/2$. If it is a , then \bar{V}_2 and V_1 have the same number of a 's. We can now apply the same argument as in step 1 and conclude that the number number of occurrences of any letter b in V_1' and in V_2' is at most 1.

Combining the above steps, we conclude that W' is balanced. \square

2.9 Extensions of the original problem

So far we have only analyzed the case where all the rates add up to one. The different results tend to prove that very few rates are balanced.

Now let us look at a generalization when all the rates do not add up to one. Assume that S is a sequence on the alphabet $\{a_1, a_2, \dots, a_k, *\}$. We only require that S is balanced for the letters a_1, \dots, a_k , but not for the special letter $*$.

On a more practical point of view, the question can be viewed as whether this allows more possibilities for rates to be balanced when “losses” are allowed (represented by the letter $*$). Then again, in general, the rates are not balanced, even if the total sum is very small as illustrated by the following lemma.

Lemma 2.28. *For an arbitrary $\varepsilon > 0$, there exists two real numbers p_1 and p_2 such that $p_1 + p_2 < \varepsilon$ and no sequence S on the alphabet $\{a, b, *\}$ with asymptotic rate p_1 for letter a and p_2 for letter b which is balanced for a and b .*

Proof. Choose two irrational numbers p_1 and p_2 with $p_1 + p_2 < \varepsilon$ such that p_1, p_2 and 1 are not linearly dependent on \mathbb{Z} . Now assume that there exists a sequence S on $\{a, b, *\}$ with asymptotic rate p_1 for letter a and p_2 for letter b which is balanced for a and b . By Theorem 2.3, then there exists two real numbers x, y such that $\delta_a(S)(n) = 1$ if $x + p_1 n \bmod 1 \in [1 - p_1, 1]$ and 0 otherwise. $\delta_b(S)(n) = 1$ if $y + p_2 n \bmod 1 \in [1 - p_2, 1]$ and 0 otherwise. In the cube $[0, 1]^2$, the set of points $(x, y) + n(p_1, p_2) \bmod (1, 1)$ is dense (see for example, Weyl's ergodic theorem [14]) and therefore hits the rectangle $[(1 - p_1, 1 - p_2), (1, 1)]$. This is not possible. \square

More on this kind of problems can be found in [16].

To end this short overview on balanced words, we must mention on the positive side that “usual” rates, such as $(1/k, 1/k, \dots, 1/k)$ are often balanced. In Appendix 4.3, some examples of balanced sequences and rates are given.

3 Routing of customers in multiple queues

The notion of balanced sequences gives birth to an large set of elegant properties in word combinatorics. However, they are rarely used in other domains. The balanced sequences in dimension $N = 2$ have been used for computer discrete line drawing [13] and also for scheduling optimization in [9, 8].

However, the case $N = 2$ behaves differently than for higher dimension (as illustrated by theorem 2.17) and does not really grasp all the complexity of the model.

Here, we present an application of balanced sequences in any dimension to scheduling optimization.

We consider a system where a sequence of tasks have to be executed by several processing units. The tasks arrive sequentially and each task can be processed by any server. The *routing* control consists in assigning to each task a server on which it will be processed. The routing is optimal if it minimizes some cost function that measures the performance of the system.

These kinds of models have been used to study load balancing within several processors in parallel processing problems as well as for efficient network utilization in telecommunication systems.

3.1 Presentation of the model

In this section we consider a more precise queueing model of the system that we described. Customers enter a multiple queue system composed of K nodes. Each node is made of several queues which form an *event graph*. Event graphs are a subset of Petri nets with no more than one input transition and one output transition per place. More details on the dynamics of event graphs can be found in [4]. In particular, their dynamical behavior is linear in the so-called $(\max, +)$ algebra ([4]). Many queueing networks can be represented as event graphs, as long as there is no inside choice for the route followed by the customers. (see [1] for a more complete discussion on this issue).

The routing of customers to the different nodes is controlled by a sequence of vectors $\{a_n\}$, with a_n is in $\{0, 1\}^K$ and $a_n^i = 1$ means that the n -th customer is routed to queue i . Note that a is a *feasible* admission sequence as long as for all n , $\sum_i a_n^i = 1$.

The link between a feasible routing policy and an infinite sequence on a finite alphabet comes from choosing the alphabet \mathcal{A} composed by the letters

$$\{(1, 0, \dots, 0), (0, 1, 0, \dots), \dots, (0, \dots, 0, 1)\}.$$

Using this alphabet on K letters, a feasible routing policy can be viewed as an infinite sequence on \mathcal{A} .

Figure 1 shows an illustration of the system we are considering.

We denote by T_n the epoch when the n -th customer enters the system. We assume that $T_1 = 0$. The inter-arrival time sequence is $\{\delta_k\} \stackrel{\text{def}}{=} \{T_{i+1} - T_i\}$. Finally, $\sigma_n^{i,j}$ will denote the service time of the n -th customer entering the j -th queue in node i .

The sequences $\{\delta_k\}$ and $\{\sigma_k^{i,j}\}$ will be considered as random processes. We also make stochastic assumption on these sequences. The inter-arrival time of the customers and the service times form stationary processes independent of each other.

3.2 Optimal admission sequence

In each node i we pick an arbitrary server s_i (which may be the last server in the node for example). The performance criteria for node i will be the traveling time to server s_i

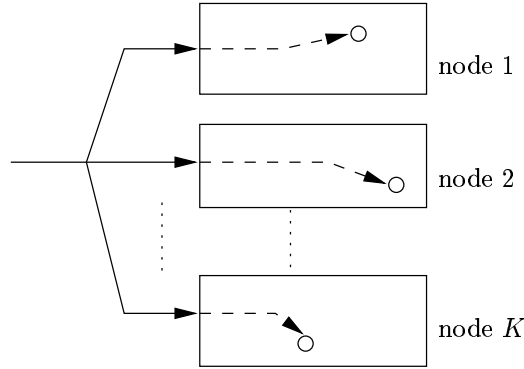


Figure 1: Illustration of the routing of customers in a K node system

of a virtual customer that would enter node i at time T_n . Under the routing policy a , this quantity only depends on the values of the n first routing choices. From the routing sequence a , we can isolate the routing decision for node i : if $a_n^i = 1$ then the customer is admitted in node i and if $a_n^i = 0$ then the customer is rejected (for node i). We denote the traveling time at time T_n by $W_n^i(a_1^i, \dots, a_n^i)$. We will be more particularly interested in the expected value of the traveling time with respect to the service times in all the servers contained in node i and with respect to the inter-arrival times: $\mathbf{W}_n^i(a_1^i, \dots, a_n^i) \stackrel{\text{def}}{=} \mathbf{E}_{\sigma, \tau} f(W_n^i(a_1^i, \dots, a_n^i))$, where $f: \mathbb{R} \rightarrow \mathbb{R}$ is any convex increasing function.

If we focus on a single node i , the function $\mathbf{W}_n^i(a_1^i, \dots, a_n^i)$ has been studied in [1]. Its most remarkable property is the fact that this function is *multimodular*. See [9, 2] for a precise definition and several properties of multimodular functions. Here, we merely point out that multimodularity is a kind of integer analog to convexity.

Proposition 3.1 ([1]). *Under the foregoing assumptions, the function $\mathbf{W}_n^i(a_1^i, \dots, a_n^i)$ satisfies the following properties.*

- (1)- $\mathbf{W}_n^i(a_1^i, \dots, a_n^i)$ is increasing in a_i ,
- (2)- $\mathbf{W}_m^i(0, \dots, 0, a_1^i, \dots, a_n^i) = \mathbf{W}_n^i(a_1^i, \dots, a_n^i)$, if $m > n$,
- (3)- $\mathbf{W}_n^i(a_1^i, \dots, a_n^i)$ is multimodular.

Proof. These properties are shown in [1]. □

Using these properties, one can derive as in [2], a lower bound $B(N, \alpha, p)$ (which is increasing in N , α and p and continuous) for any routing a^i , for the following discounted cost.

$$\sum_{n=1}^{\infty} (1 - \alpha) \alpha^{n-1} \mathbf{W}_n^i(a_1^i, a_2^i, \dots, a_n^i) \geq B(N, \alpha, p_\alpha), \quad (5)$$

where $p_\alpha \stackrel{\text{def}}{=} (1 - \alpha) \sum_{k=1}^{\infty} \alpha^{k-1} a_k$.

Also, for a given p , we define the regular sequence with rate p and an arbitrary phase θ ,

$$a^p(\theta) \stackrel{\text{def}}{=} \lfloor np + \theta \rfloor - \lfloor (n-1)p + \theta \rfloor,$$

(see the definition 2.2). One can show as in [2] that $a^p(\theta)$ satisfies

$$\lim_{m \rightarrow \infty} \frac{1}{m} \sum_{n=1}^m \mathbf{W}_n^i(a_1^p(\theta), \dots, a_n^p(\theta)) = B_i(N, 1, p). \quad (6)$$

Here, however, we are interested in the performance of all nodes together. Therefore, we choose as a cost function, the undiscounted average on n of some linear combination of the expected traveling time in all nodes.

Let h be any increasing linear function, $h : \mathbb{R}^K \rightarrow \mathbb{R}$. We consider the undiscounted average cost of a feasible routing sequence a ,

$$g(a) \stackrel{\text{def}}{=} \overline{\lim}_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N h(\mathbf{W}_n^1, \dots, \mathbf{W}_n^K).$$

From this point on, we mimic the general method developed in [2] for our case. We use the following notations.

$$B_i(\alpha, p_i) \stackrel{\text{def}}{=} \sup_N B_i(N, \alpha, p_i),$$

and

$$B_i(p_i) \stackrel{\text{def}}{=} \sup_{\alpha \leq 1} B_i(\alpha, p_i).$$

Our objective is to minimize $g(a)$.

Theorem 3.2. *The following lower bound holds for all policies:*

$$g(a) \geq \inf_{p_1 + \dots + p_K = 1} h(B_1(p_1), \dots, B_K(p_K)).$$

Proof. Due to Littlewood's and Jensen's inequalities as well as Equation 5, we have

$$\begin{aligned} & \overline{\lim}_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N h(\mathbf{W}_n^1, \dots, \mathbf{W}_n^K) \\ & \geq \overline{\lim}_{\alpha \rightarrow 1} (1 - \alpha) \sum_{n=1}^{\infty} \alpha^{n-1} h(\mathbf{W}_n^1, \dots, \mathbf{W}_n^K) \\ & \geq \overline{\lim}_{\alpha \rightarrow 1} h \left((1 - \alpha) \sum_{n=1}^{\infty} \alpha^{n-1} \mathbf{W}_n^1, \dots, (1 - \alpha) \sum_{n=1}^{\infty} \alpha^{n-1} \mathbf{W}_n^K \right) \\ & \geq \overline{\lim}_{\alpha \rightarrow 1} h(B_1(\alpha, p_1^a(\alpha)), \dots, B_K(\alpha, p_K^a(\alpha))) \end{aligned} \quad (7)$$

where

$$p_i^a(\alpha) \stackrel{\text{def}}{=} (1 - \alpha) \sum_{k=1}^{\infty} \alpha^{k-1} a_k^i. \quad (8)$$

We note that $\sum_{i=1}^K p_i^a(\alpha) = 1$. Hence, one may choose a sequence $\alpha_n \uparrow 1$ such that the following limits exist:

$$\lim_{n \rightarrow \infty} p_i^a(\alpha_n) = p_i, \quad i = 1, \dots, K \quad (9)$$

and $\sum_{i=1}^K p_i = 1$. From the continuity of $B_i(\alpha, p_i)$ in p and α we get from (7)

$$\begin{aligned} g(a) &\geq h(B_1(p_1), \dots, B_K(p_K)) \\ &\geq \inf_{p_1 + \dots + p_K = 1} h(B_1(p_1), \dots, B_K(p_K)). \end{aligned} \quad (10)$$

□

Note that there exists some p^* that achieves the infimum

$$\inf_{p_1 + \dots + p_K = 1} h(B_1(p_1), \dots, B_K(p_K)),$$

since $h(B_1(p_1), \dots, B_K(p_K))$ is continuous in $p = (p_1, \dots, p_K)$.

Consider $\theta = (\theta_1, \dots, \theta_K)$ and the routing policy $a^{p^*}(\theta)$ given for each i by

$$a_k^{i,p^*}(\theta_i) = \lfloor kp_i^* + \theta_i \rfloor - \lfloor (k-1)p_i^* + \theta_i \rfloor. \quad (11)$$

There are some p^* for which the condition of feasibility of the policy $a^{p^*}(\theta)$ is satisfied, that is, there exists some $\theta = (\theta_1, \dots, \theta_K)$, such that the regular policy $a^{p^*}(\theta)$ is feasible.

Using the correspondence between a routing policy and a sequence on the alphabet \mathcal{A} , these p^* correspond precisely to *balanced* rates.

Theorem 3.3. *Assume that p^* is balanced. Then $a^{p^*}(\theta)$ is optimal for the average cost, i.e. it minimizes $g(a)$ over all feasible policies.*

Proof. The proof follows directly from Theorem 3.2 and Equation (6). □

4 Study of some special cases

The problem which remains to be addressed is to find in which cases, the rate vector p^* is balanced. We will present several simple examples for which we can make sure the optimal rate p^* is balanced.

4.1 The case $K = 2$

If $K = 2$, then, the optimal rate vector is of the form $p^* = (p_1^*, 1 - p_1^*)$. Theorem 2.17 says that p^* is always balanced and therefore, the optimal routing sequence is given by an associated balanced sequence. Note that this approach does not give any direct way to compute the value of p^* , however, it gives the structure of the optimal policy.

4.2 The homogeneous case

Now let K be arbitrary and each node is made of a single server, all servers being identical. This model is displayed in Figure 2.

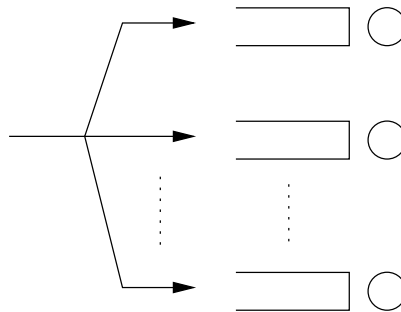


Figure 2: Routing in homogeneous queues.

Also assume that the function h is symmetrical in all coordinates (for example, just the sum of all waiting times) By symmetry and convexity in (p_1, \dots, p_K) of the function $h(V_1^1(p_1), \dots, V_K^1(p_K))$, then, $p^* = (1/K, \dots, 1/K)$, which is balanced. The associated balanced sequence is the round robin routing scheme. Applying Theorem 3.3 yields the following result which is new (to the best of the author's knowledge).

Theorem 4.1. *The round robin routing to K identical ./G/1 queues, minimizes the workload for the convex increasing order in all the queues over all admission sequences with no information on the state of the system.*

In [11], the round robin routing is proved to be optimal for a slightly different cost function for K identical ./GI/1 queues. Their method uses an intricate coupling argument, where as our proof is a simple corollary of the general theory on multimodular functions.

To illustrate the advantage of our approach, we further generalize the result to a system composed of K identical $(\max, +)$ linear systems with a single entry. In this case, the symmetry argument used in the case of simple queues still holds. Then again, the round robin routing policy minimizes the traveling time in each system. This case includes models such as routing among several identical systems composed of queues in tandem, for example (see Figure 3).

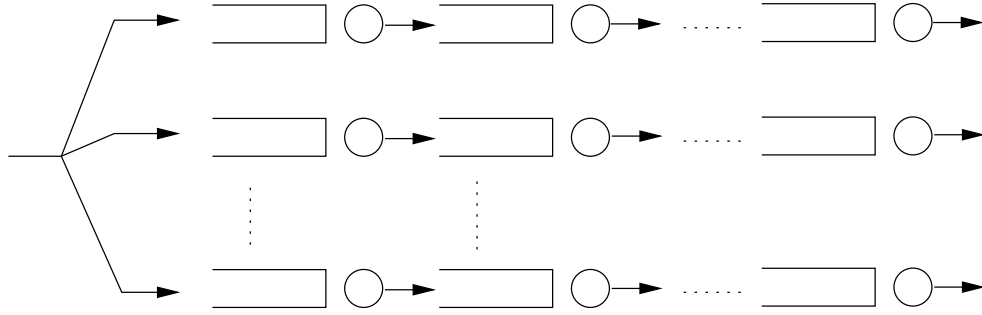


Figure 3: Routing in queues in tandem

4.3 Two sets of identical servers

As a consequence of the two previous cases, we can consider a system composed of K_1 identical queues of type 1 and K_2 queues of type 2. Again, assume that h is symmetrical. Then, by symmetry arguments, the optimal rate vector p^* is of the form

$$\left(\frac{p}{K_1}, \dots, \frac{p}{K_1}, \frac{1-p}{K_2}, \dots, \frac{1-p}{K_2}\right).$$

This sequence is balanced indeed, which implies that the optimal routing is of balanced type.

Many other examples of this kind can be derived from these examples through similar constructions.

Appendix

This is a collection of balanced rates. For some rates, we give a corresponding balanced word.

- $(1/7, 2/7, 4/7)$ is balanced and $S = (abacaba)^\omega$.
- $(1/11, 2/11, 4/11, 4/11)$ is balanced and $S = (abcababcabd)^\omega$.
- $(1/11, 2/11, 2/11, 6/11)$ is balanced and $S = (abacaabacad)^\omega$.
- $(1/11, 1/11, 3/11, 6/11)$ is balanced and $S = (acabaabadab)^\omega$.
- $(1/14, 1/14, 4/14, 8/14)$ is balanced and $S = (abacabaabadaba)^\omega$.
- For all real number $0 < p \leq 1$, the rates $(1-p, p/4, p/4, p/2)$, with a balanced word constructed from a regular word with rate p where all 1 are replaced in turn by the sequence $(abc)^\omega$ and each 0 by the letter d .

- for all N , $((2^{N-1}/(2^N - 1), \dots, 2^{N-i}/(2^N - 1), \dots, 1/(2^N - 1)))$ is balanced. The associated balanced sequence is constructed recursively. If S_{N-1} is balanced on the alphabet $\{a_2, \dots, a_N\}$, $S_{N-1} = W_{N-1}^\omega$, then $W_N = W_{N-1}a_1W_{N-1}$ and $S_N = W_N^\omega$.
- $(1/k, \dots, 1/k)$, a balanced word with those letters is: $S = (a_1a_2a_3 \dots a_k)^\omega$.
- $(p, \dots, p, \beta, \dots, \beta)$. A balanced word with those rates is constructed in the following way: Choose a balanced word S on letters, (A, B) with rate $(\sum_{i=1}^k p, \sum_{i=1}^h \beta)$. In S replace all the A (resp. B) by a_1, a_2, \dots, a_k (resp. b_1, \dots, b_h) in a round robin fashion to get a balanced word with the required rates.

Acknowledgment

The authors would like to thank Alain Jean-Marie who pointed out the reference [18] and gave them a first proof of Lemma 2.10.

References

- [1] Eitan Altman, Bruno Gaujal, and Arie Hordijk. Admission control in stochastic event graphs. Technical report, INRIA, May 1997.
- [2] Eitan Altman, Bruno Gaujal, and Arie Hordijk. Multimodularity, convexity and optimization properties. Technical report, INRIA, May 1997.
- [3] Arnoux, Mauduit, Shiokawa, and Tamura. Complexity of sequences defined by billiards in the cube. *BSMF: Bulletin de la Societ'e Mathematique de France*, 122, 1994.
- [4] F. Baccelli, G. Gohen, G.J. Olsder, and J.-P. Quadrat. *Synchronization and Linearity*. Wiley, 1992.
- [5] Y. Baryshnikov. Complexity of trajectories in rectangular billiards. *Com. Math. Phys.*, 175:43–56, 1995.
- [6] J. Berstel and P. Séébold. Morphismes de Sturm. *bullbelgmathsoc*, 1:175–189, 1994.
- [7] E.M. Coven and G.A. Hedlund. Sequences with minimal block growth. *Math. System Theory*, 7(2), 1973.
- [8] Bruno Gaujal. Optimal allocation sequences of two processes sharing a resource. *Discrete Event Dynamic Systems*, 1997. To appear.
- [9] Bruce Hajek. Extremal splittings of point processes. *Mathematics of Operation research*, 10(4), November 1985.
- [10] P. Hubert. Suites équilibrées. Technical report, Université de Marseille, 1996.

- [11] Zhen Liu and Rhonda Righter. Optimal load balancing on distributed homogeneous unreliable processors. *Journal of Operation Research*, 1997. To appear.
- [12] J. A. Loeve. *Markov Decision Chains With Partial Information*. PhD thesis, Leiden University, 1995.
- [13] M. Lothaire. *Mots*, chapter Tracé de droites, fractions continues et morphismes itérés (J. Berstel). Hermes, 1991.
- [14] Ya. G. Sinai. *Introduction to Ergodic Theory*. Princeton University Press, Princeton, N.J., 1976.
- [15] R. Tijdeman. On complementary triples of sturmian sequences. Technical Report W95-13, Leiden University, The Netherlands, November 1995.
- [16] R. Tijdeman. On disjoint pairs of sturmian bisequences. Technical report, Leiden University, The Netherlands, 1995.
- [17] L. Vuillon. Combinatoire des motifs d'une suite sturmienne bidimensionnelle. *Theoretical Computer Science*. To appear.
- [18] Herbert S. Wilf. *Generatingfunctionology*. Academic Press, 2nd edition, 1994.



Unité de recherche INRIA Lorraine, Technopôle de Nancy-Brabois, Campus scientifique,
615 rue du Jardin Botanique, BP 101, 54600 VILLERS LÈS NANCY
Unité de recherche INRIA Rennes, Irisa, Campus universitaire de Beaulieu, 35042 RENNES Cedex
Unité de recherche INRIA Rhône-Alpes, 655, avenue de l'Europe, 38330 MONTBONNOT ST MARTIN
Unité de recherche INRIA Rocquencourt, Domaine de Voluceau, Rocquencourt, BP 105, 78153 LE CHESNAY Cedex
Unité de recherche INRIA Sophia Antipolis, 2004 route des Lucioles, BP 93, 06902 SOPHIA ANTIPOLIS Cedex

Éditeur
INRIA, Domaine de Voluceau, Rocquencourt, BP 105, 78153 LE CHESNAY Cedex (France)
ISSN 0249-6399