Working Set Size Strings

Domenico Ferrari

Computer Science Division
Department of Electrical Engineering and Computer Sciences
and the Electronics Research Laboratory
University of California, Berkeley

ABSTRACT

The properties of working set size strings, i.e., of the strings of integers representing the sequences of values taken by the working set sizes of programs, are investigated. Particular attention is given to the conditions to be satisfied by a given string of integers, by a pair of such strings, or by an n-tuple of strings for the existence of a page reference string whose working set size strings coincide with the given ones for given values of the window size. These conditions are useful in generating artificial reference strings exhibiting a given memory demand dynamics.

1. Introduction

The working set policy for memory management in virtual-memory systems was proposed by Denning in 1968 [1] and has been extensively studied since then (see for example [2]). Only recently, however, has a generative model of program behavior based on the working set concept been introduced [3]. This model is intended to be the working-set-oriented counterpart of the well-known LRU stack model, that was inspired by the LRU replacement policy [4]. The original proposal of the new model presented a deterministic approach, but a stochastic version [5] was soon added to it. The new model is able to produce a reference string which followed a given memory demand pattern in virtual time; this pattern is represented by the dynamics of the size of the working set. The stochastic version allows the model's user to describe this dynamic behavior by assigning the values of a limited number of parameters rather than having to specify the salient points of the working set size curve.

While investigating the viability of the new model, it was soon discovered that a reference string corresponding to a given working set size dynamics does not always exist. This observation obviously raised the question of what conditions must be satisfied by a given dynamics for the existence of a reference string whose dynamics coincides with the given one. The study of this problem inevitably led to the study of the properties of working set size strings, that are the subject of this paper.

We shall begin by introducing some basic definitions.

The research reported here has been supported in part by the National Science Foundation under Grant MCS80-12900.

Definition 1. Given a set P, a *reference string* over P is a finite string $r = r_1, r_2, \ldots r_n$, with $r_t \in P$ for $t = 1, 2, \ldots n$.

Definition 2. Given a reference string r of length n and a positive integer T called the *window size*, the *working set* of r at t, with window size T and $1 \le t \le n$, is the set

$$W(r, t, T) = \{r_j \mid t_0 \le j \le t, t_0 = \max(t - T + 1, 1)\} .$$

The size of the working set at the time t of a given reference string r with window size T,

$$w(r, t, T) = | W(r, t, T) | ,$$

is a positive integer bounded by T and $|P|$:

$$1 \le w(r, t, T) \le \min(T, |P|), \quad (t = 1, 2, \ldots n) .$$

When it will be desirable to define the value of the working set size for $t = 0$ or for $t > n$, we shall set, for all r and T, $w(r, 0, T) = 0$, and $w(r, t, T) = w(r, n, T)$ for $t > n$.

Definition 3. To any reference string r of length n and window size T there corresponds a string $w = w_1, w_2, \ldots w_n$ called the *working set size (wss) string* of r for window size T, with $w_t = w(r, t, T)$, $(t = 1, 2, \ldots n)$.

The string of working set sizes does not fully represent the dynamic characteristics of a reference string. For example, $w_t = w_{t-1}$ does not always mean that $r_t \in W(r, t-1, T)$: the equality holds also when $r_t \notin W(r, t-1, T)$ and $r_{t-T} \notin W(r, t, T)$. In other words, the size remains unchanged not only when the next reference is to a page already in the working set, but also when one of the members of $W(r, t-1, T)$ drops out and a new member joins the working set at t. Since these two cases correspond to two very different dynamic behaviors of the working set, it is useful to distinguish them by introducing the notion of flat fault.

Definition 4. Given a reference string r and a window size T, a *flat fault* is said to occur at t if

$$r_t \notin W(r, t - 1, T) \text{ and } r_{t-T} \notin W(r, t, T) .$$

Definition 5. Given a reference string r of length n and a window size T, the *flat fault (ff) string* associated to r and T is the Boolean string

$$f(r, t, T) = \begin{cases} 1 \text{ if } r \text{ has a flat fault at } t \\ 0 \text{ otherwise} \end{cases}$$

for $t = 1, 2, \ldots n$.

When needed, we shall assume $f(r, t, T) = 0$ also for $t > n$.

Definition 6. A string of integers w and a Boolean string f are said to be *compatible* if they have the same length n and, for all t $(1 \le t \le n)$ such that $f_t = 1$, it is $w_t = w_{t-1}$.

To any given reference string r and window size T there corresponds a pair (w, f) of compatible strings, the wss string and the ff string.

2. Single-window properties

In this section, an answer will be provided to the question of when is a given string of positive integers the wss string of a real reference string for a given window size.

Definition 7. A finite string $w = w_1, w_2, \ldots w_n$ of positive integers is a *feasible working set size (fwss) string* for window size T if there exists at least one reference string r of length n that has w as its working set size string for window size T. Similarly, a pair (w, f) of compatible strings of length n is a *feasible pair* for window size T if there exists at least one reference string r of length n that has w as its working set size string and f as its flat fault string for window size T.

Definition 8. Given a string $w = w_1, w_2, \ldots w_n$ of positive integers, the *decrement string* $d = d_1, d_2, \ldots d_n$ associated to w is the string of binary digits defined as

$$d_1 = 0$$

$$d_t = \begin{cases} 1 \text{ if } w_t < w_{t-1} \text{ or } f_t = 1 \\ 0 \text{ otherwise} \end{cases} \quad (t = 2, 3, \ldots n) .$$

Note that, in Definition 7, a flat fault is treated as a decrement. Indeed, a flat fault may be seen as the superposition in time of a decrement and of an increment in the working set size. The algorithm on which the proof of part (2) of Theorem 1 below is based looks at flat faults from both viewpoints.

Under what conditions is a given string w of positive integers an fwss string?

Theorem 1. A pair (w, f) of compatible strings of length n is a feasible pair for window size T if and only if

(i) $w_1 = 1$;

(ii) $|w_t - w_{t-1}| \leq 1$ for $t = 2, 3, \ldots n$;

(iii) $w_t \leq T$ for $t = 1, 2, \ldots n$;

(iv) $\sum_{i=1}^{T-1} d_{t+i} < w_t$ for $t = 1, 2, \ldots n - T + 1$.

Conditions (i) – (iv) will be referred to in the sequel as the *fwss conditions*.

Proof. A full proof of this theorem can be found in [6]. Here, only a sketch of the proof will be given.

(1) The fwss conditions are necessary. The necessity of conditions (i), (ii), and (iii) stems from Definition 2. That of condition (iv) is proved by observing that the maximum number of members of P that may drop out of the working set between $t + 1$ and $t + T - 1$ is $w_t - 1$. Indeed, r_t does not drop out

since $r_t \in W(r, t + T - 1, T)$; also, none of the members of F that join the working set during the interval being considered may drop out before $t + T + 1$.

(2) The fwss conditions are sufficient. This part of the proof is constructive: a string generation algorithm is shown to be able to produce a reference string with the given working set size dynamics and to run to completion if the given w satisfies the fwss conditions. The algorithm maintains three disjoint and exhaustive subsets of P: the *candidate set* C, the *forbidden set* F, and the *external set* E. For all t, these sets satisfy the relationships

$$C_t \cup F_t = W(r, t, T), \quad E_t \cup W(r, t, T) = P,$$
$$C_t \cap F_t \cap E_t = \phi.$$

Initially, at $t = 0$, sets C_0 and F_0 are empty, and $E_0 = P$. At time t $(t = 1, 2, \dots n)$, the following two cases may occur:

(a) $w_t > w_{t-1}$ or $f_t = 1$; then, $r_t \in E_{t-1}$;
(b) $w_t \le w_{t-1}$; then, $r_t \in C_{t-1}$.

The three sets are updated, after generating the next reference r_t at time t, as follows:

(c) if $w_{t+T} \ge w_{t+T-1}$ and $f_{t+T} = 0$, then r_t remains in C_t (if it was in C_{t-1}) or is moved from E_{t-1} into C_t, and is given a deadline equal to $t + T$; this means that it will have to be referenced again not later than time $t + T$;
(d) if $w_{t+T} < w_{t+T-1}$ or $f_{t+T} = 1$, then r_t is moved from either E_{t-1} or C_{t-1} into F_t, and is given a deadline equal to $t + T$; this means that it cannot be referenced again until after $t + T$, at which time it will automatically move from F_{t+T-1} into E_{t+T}.

Even though it is not necessary, assuming that the three sets (especially C) are dealt with as FIFO queues simplifies the proofs.

The only problems that could arise during the generation of the reference string are the following:

(a) $E_{t-1} = \phi$: this cannot happen when case (a) occurs if fwss condition (*iii*) is satisfied;
(b) $C_{t-1} = \phi$: this cannot happen when case (b) occurs if fwss conditions (*ii*) and (*iv*) are satisfied;
(c) the deadline of an element of C_t is t: this cannot happen if fwss conditions (*ii*) and (*iii*) are satisfied.

■

Note that a special case of Theorem 1 is the one in which $f_t = 0$ for all t. Thus, the fwss-conditions are also the necessary and sufficient conditions for the existence of a reference string having a given wss string and no flat faults [3].

3. Two-window properties

In this section, the generation of an artificial reference string that follows the memory demand patterns specified by its wss string for two different window sizes is discussed. This approach to artificial string generation has been suggested by the need to improve the accuracy of the generated strings in non-working-set (e.g., LRU) environments [7] [8].

Definition 9. Given a feasible pair (w^1, f^1) of strings of length n for window size T^1, and a feasible pair (w^2, f^2) of strings of length n for window size T^2, (w^1, f^1) and (w^2, f^2) are said to be *consistent* if there exists a reference string r of length n such that

$$w(r, t, T^1) = w_t^1, \quad f(r, t, T^1) = f_t^1,$$
$$w(r, t, T^2) = w_t^2, \quad f(r, t, T^2) = f_t^2, \qquad \text{for } t = 1, 2, \ldots n .$$

Theorem 2. Two pairs of strings (w^1, f^1) and (w^2, f^2), that are feasible for window sizes T^1 and T^2, respectively, with $T^2 - T^1 = \Delta > 0$, are consistent if and only if:

(a) $w_t^2 \geq w_t^1$, $(t = 1, 2, \ldots n)$;

(b) for any t such that $w_t^2 > w_{t-1}^2$ or $f_t^2 = 1$, it is $w_t^1 > w_{t-1}^1$ or $f_t^1 = 1$;

(c) for any t such that $d_t^2 = 1$, it is $d_{t-\Delta}^1 = 1$;

(d) for any t such that $w_t^2 \leq w_{t-1}^2$ (with $f_t^2 = 0$) and either $w_t^1 > w_{t-1}^1$ or $f_t^1 = 1$, there exists an index value, say $t - K$ (with K integer and $0 < K < \Delta$), such that $d_{t-K}^1 = 1$ and $d_{t-K+\Delta}^2 = 0$. The correspondence between index values $t - K$ and t is one-one; in other words, each t at which the situation described in this condition occurs must have its exclusive corresponding index value $t - K$ satisfying the condition.

Conditions (a) - (d) will be referred to in the sequel as the *consistency conditions*.

Proof. (1) The conditions are necessary. Let there be a reference string r of which w^1, w^2 are the wss strings, and f^1, f^2 are the ff strings for window size T^1 and T^2, respectively.

(a) The necessity of this condition is a direct consequence of the definition of working set (Definition 2): for all t, we have $W_t^2 \supseteq W_t^1$. Note that W_t^2 is a short-hand symbol for $W(r, t, T^2)$.

(b) Since $r_t \notin W_{t-1}^2$, we also have $r_t \notin W_{t-1}^1$, as $W_i^2 \supseteq W_i^1$ for all i.

(c) Since $d_t^2 = 1$, we have $r_{t-T^2} \notin W_t^2$, and, since $W_{t-\Delta}^1 \subseteq W_t^2$, it must be $r_{t-T^2} \notin W_{t-\Delta}^1$. Hence, $d_{t-\Delta}^1 = 1$.

(d) Let r_t be a reference to page $p_i \in P$. Since we have $p_i \in W_{t-1}^2$ and $p_i \notin W_{t-1}^1$, there is at least one reference to p_i between $t - T^2 + 1$ and $t - T^1 - 1$. Let $t - T^1 - K$, with $0 < K < \Delta$, be the time prior to t of the most recent reference to p_i. Then, $p_i \in W_{t-K-1}^1$ and $p_i \notin W_{t-K}^1$, hence $d_{t-K}^1 = 1$. Also, $p_i \in W_{t-K+\Delta-1}^2$ and, since $r_t = p_i$, $p_i \in W_{t-K+\Delta}^2$; hence, $d_{t-K}^2 = 0$.

(2) The conditions are sufficient. Let (a) through (d) be satisfied. The algorithm described in part (2) of the proof of Theorem 1 can be easily extended to the case of two window sizes: the tree sets C, F, and E will have to be defined and updated for each of the two values of T, and the next reference of the string being generated will be a member of the intersection of the appropriate sets. For instance, if both $w_t^2 > w_{t-1}^2$ and $w_t^1 > w_{t-1}^1$, r_t will be chosen from $E_{t-1}^2 \cap E_{t-1}^1$.

Since condition (b) is satisfied, we cannot have $w_t^2 > w_{t-1}^2$ (or $f_t^2 = 1$) and, at the same time, $w_t^1 \le w_{t-1}^1$ and $f_t^1 = 0$. The cases which may occur at time t can be grouped into the following three categories.

(i) $w_t^2 > w_{t-1}^2$ (or $f_t^2 = 1$) and $w_t^1 > w_{t-1}^1$ (or $f_t^1 = 1$). In this case, we must choose $r_t \in E_{t-1}^2 \cap E_{t-1}^1$. Since w^2 is to receive a new element of P, it must be $E_{t-1}^2 \ne \phi$. Because of condition (a), we have $W_i^1 \subseteq W_i^2$, hence $E_i^1 \supseteq E_i^2$ for all i. Thus, $E_{t-1}^1 \supseteq E_{t-1}^2$, and $E_{t-1}^2 \cap E_{t-1}^1 \ne \phi$.

(ii) $w_t^2 \le w_{t-1}^2$ and either $w_t^1 > w_{t-1}^1$ or $f_t^1 = 1$. In this case, $r_t \in C_{t-1}^2 \cap E_{t-1}^1$. By condition (d), there exists an integer K $(0 < K < \Delta)$ such that $d_{t-K}^1 = 1$. Let r_{t-K-T^1} be a reference to page p_j. At $t - K - T^1$, immediately after being referenced, p_j joins F^1 and, due to condition (d), C^2. At time $t - K$, p_j moves from F^1 to E^1, but remains in C^2. At time t, p_j is referenced again for the first time since $t - K - T^1$ (this is how K is defined according to condition (d)), and therefore can still be found in C_{t-1}^2 and E_{t-1}^1. Hence, $C_{t-1}^2 \cap E_{t-1} \ne \phi$.

(iii) $w_t^2 \le w_{t-1}^2$ and $w_t^1 \le w_{t-1}^1$. In this case, we must choose $r_t \in C_{t-1}^2 \cap C_{t-1}^1$. If $w_t^1 = w_{t-1}^1$, and r_{t-T^1} is a reference to page p_j, then $p_j \in C_{t-1}^1$. Since it is also $p_j \in C_{t-1}^2$, we have $C_{t-1}^2 \cap C_{t-1}^1 \ne \phi$. If, on the other hand, $w_t^1 < w_{t-1}^1$, it is possible that $w_{t+1}^1 \ge w_t^1$, and hence r_{t-T^1+1} will be in C_{t-1}^1. C_{t-1}^1 will be empty only when either $w_i^1 < w_{i-1}^1$ or $f_i^1 = 1$ for $i = t, t+1, \ldots t + T^1 - 1$. However, this is impossible since w^1 is an fwss string, and satisfies fwss condition (iv). Note that, if $p_j \in C_{t-1}^1$, then, by condition (c), $p_j \in C_{t-1}^2$. Hence, $C_{t-1}^2 \cap C_{t-1}^1 \ne \phi$.

∎

Note that, unlike what happens with a single window size characterization, the consistency conditions for the flat fault free case may not be satisfied by the two pairs (w^1, f^1) and (w^2, f^2) characterizing a real reference string if this string has flat faults [9].

4. Multiple-window properties

Let us now consider m feasible pairs (w^1, f^1), (w^2, f^2), ... (w^m, f^m) of strings of length n, whose window sizes are all different. Without loss of generality, we can assume that $T^1 < T^2 < \ldots < T^m$.

Lemma 1. Given m feasible pairs (w^1, f^1), (w^2, f^2), ... (w^m, f^m) of strings of length n, with $T^1 < T^2 < \ldots < T^m$, if (w^1, f^1) and (w^2, f^2) are consistent, (w^2, f^2) and (w^3, f^3) are consistent,... and (w^{m-1}, f^{m-1}) and (w^m, f^m) are consistent, then (w^1, f^1) and (w^m, f^m) are consistent.

Proof. (a) Since $w_t^1 \le w_t^2$, $w_t^2 \le w_t^3$, ... and $w_t^{m-1} \le w_t^m$, we have $w_t^1 \le w_t^m$. Hence, w^1 and w^m satisfy consistency condition (a).

(b) By consistency condition (b), if either $w_t^m > w_{t-1}^m$ or $f_t^m = 1$, then either $w_t^{m-1} > w_{t-1}^{m-1}$ or $f_t^{m-1} = 1$. Also, if either $w_t^{m-1} > w_{t-1}^{m-1}$ or $f_t^{m-1} = 1$, then either $w_t^{m-2} > w_{t-1}^{m-2}$ or $f_t^{m-2} = 1$, and so on. Hence, w^1 and w^m satisfy consistency condition (b).

(c) Let $\Delta_i = T^{i+1} - T^i$ for $i = 1, 2, \ldots m - 1$. By definition, $\Delta_i > 0$ for all i.

If $d_t^m = 1$, then $d^{m-1}_{t-\Delta_{m-1}} = 1$. Also, if $d^{m-1}_{t-\Delta_{m-1}} = 1$, then $d^{m-2}_{t-\Delta_{m-1}-\Delta_{m-2}} = 1$, and so on, until we find $d^1_{t-\sum_{i=1}^{m-1}\Delta_i} = 1$. Since $\sum_{i=1}^{m-1}\Delta_i = T^m - T^1$, consistency condition (c) is satisfied by w^m and w^1.

(d) By consistency condition (d), whenever $w_t^m \le w_{t-1}^m$ and either $w_t^{m-1} > w_{t-1}^{m-1}$ or $f_t^{m-1} = 1$, there exists an index value $t - K$ (with $0 < K < \Delta_{m-1}$) such that $d^{m-1}_{t-K} = 1$ and $d^m_{t-K+\Delta_{m-1}} = 0$. By consistency condition (c), since $d^{m-1}_{t-K} = 1$, it is also $d^{m-2}_{t-K+\Delta_{m-2}} = 1$. By consistency condition (b), since either $w_t^{m-1} > w_{t-1}^{m-1}$ or $f_t^{m-1} = 1$, we have either $w_t^{m-2} > w_{t-1}^{m-2}$ or $f_t^{m-2} = 1$. Hence, condition (d) is satisfied by strings w^m and w^{m-2}. By repeatedly applying the same arguments to $w^{m-3}, \ldots w^2, w^1$, we can prove that w^m and w^1 satisfy consistency condition (d). ∎

Note that the transitivity of pair consistency is generally restricted to that described in Lemma 1. For example, if (w^3, f^3) and (w^2, f^2) are consistent, and (w^2, f^2) and (w^1, f^1) are consistent, (w^3, f^3) and (w^1, f^1) are not guaranteed to be consistent.

We now generalize Definition 9 to the case of m window sizes.

Definition 10. Given pairs (w^1, f^1), (w^2, f^2), ... (w^m, f^m) of strings, each of length n, that are feasible for window sizes $T^1, T^2, \ldots T^m$, respectively, these pairs are said to be *consistent* if there exists a reference string r such that, for $t = 1, 2, \ldots n$,

$$
\begin{aligned}
w(r, t, T^1) &= w_t^1, & f(r, t, T^1) &= f_t^1, \\
w(r, t, T^2) &= w_t^2, & f(r, t, T^2) &= f_t^2, \\
&\cdots\cdots & &\cdots\cdots \\
w(r, t, T^m) &= w_t^m, & f(r, t, T^m) &= f_t^m.
\end{aligned}
$$

Under what conditions are m given feasible pairs all consistent? The answer to this question is provided by the following theorem.

Theorem 3. Given m feasible pairs (w^1, f^1), (w^2, f^2), ... (w^m, f^m) of strings of length n, with $T^1 < T^2 < \ldots < T^m$, if (w^i, f^i), and (w^{i+1}, f^{i+1}) are consistent for $i = 1, 2, \ldots m-1$, then all pairs are consistent.

Proof. The approach is the one followed in part (2) of the proof of Theorem 2. Let all strings be pairwise consistent. The algorithm for string generation can be easily extended to the case of m window sizes. For each window size there will be three sets C, F, and E, that will be updated independently after the generation of each reference. The reference at time t will be a member of the intersection of the appropriate sets defined at $t-1$. Because of the consistency conditions that are satisfied by all pairs of strings with adjacent apex values, at time t we can have one of the following three cases.

(i) Either $w_t^i > w_{t-1}^i$ or $f_t^i = 1$ for all $i = 1, 2, \ldots m$. In this case, $r_t \in \bigcap_{i=1}^{m} E_{t-1}^i$. Since $w_t^m > w_{t-1}^m$, we must have $E_{t-1}^m \neq \phi$. Also, since by consistency condition (a) $W_{t-1}^j \subseteq W_{t-1}^{j+1}$ for $j = 1, 2, \ldots m-1$, it is

$E^j_{t-1} \supseteq E^{j+1}_{t-1}$. Hence, $\bigcap_{i=1}^{m} E^i_{t-1} \neq \phi$.

(ii) Either $w^i_t > w^i_{t-1}$ or $f^i_t = 1$ for $i = 1, 2, ... v$, and $w^j_t \leq w^j_{t-1}$ for $j = v + 1, v + 2, ... m$. In this case, $r_t \in \left[\left[\bigcap_{i=1}^{v} E^i_{t-1} \right] \cap \left[\bigcap_{j=v+1}^{m} C^j_{t-1} \right] \right]$. By Lemma 1, pairs (w^m, f^m) and (w^v, f^v) are consistent, and so are (w^m, f^m) and (w^{v-1}, f^{v-1}) , (w^m, f^m) and (w^{v-2}, f^{v-2}) ,... (w^m, f^m) and (w^1, f^1) . All these pairs of pairs satisfy consistency condition (d), according to which only an element of P can be referenced at t, namely, the one that was last referenced at $t - K - T^1$, with $0 < K < T^m - T^1$ (see the proof of Theorem 2, part (2) (ii)). Applying condition (d) to any pair of string pairs composed of one pair with window size smaller than T^{v+1} and one pair with window size larger than T^v will result in the same subset of elements of P that can be referenced at t. This subset is given by $\left[\bigcap_{i=1}^{v} E^i_{t-1} \right] \cap \left[\bigcap_{j=v+1}^{m} C^j_{t-1} \right]$, and is guaranteed by condition (d) not to be empty.

(iii) $w^i_t \leq w^i_{t-1}$ for all $i = 1, 2, ... m$. In this case, $r_t \in \bigcap_{i=1}^{m} C^i_{t-1}$. If $w^1_t \leq w^1_{t-1}$, and r_{t-T} is a reference to page p_j , then $p_j \in C^1_{t-1}$. Because of the consistency of (w^1, f^1) and (w^h, f^h) $(h = 2, 3, ... m)$, condition (c) yields $p_j \in C^h_{t-1}$. Thus, $\bigcap_{i=1}^{m} C^i_{t-1} \neq \phi$. If, on the other hand, $w^1_t < w^1_{t-1}$, by Theorem 1 there must be an integer l, with $0 < l < T^1$, such that $w^1_{t+l} \geq w^1_{t+l-1}$. Then, by consistency condition (c), $w^h_{t+l+T^h-T^1} \geq w^h_{t+l+T^h-T^1-1}$ for $h = 2, 3, ... m$. This means that at least the element of P referenced at time $t + l - T^1$ is a member of all C^i_{t-1}'s. Thus, $\bigcap_{i=1}^{m} C^i_{t-1} \neq \phi$.

■

Note that Theorem 3 has proved that the consistency of $m - 1$ ordered pairs of (w, f) pairs is sufficient to guarantee that all pairs will be consistent with each other. It is obvious, however, that this condition is also necessary.

5. Summary

Some of the properties of working set size strings useful in the generation of artificial page reference strings exhibiting a given dynamics of memory demand have been investigated. Necessary and sufficient conditions have been found for the existence of a reference string having

(a) a given working set size string (or wss string and ff string) for a given window size (Theorem 1);

(b) a given pair of working set size strings (or (w, f) pairs) for given window sizes T^1 and T^2 (Theorem 2);

(c) a given m-tuple of working set size strings (or (w, f) pairs) for m given window sizes $T^1, T^2, ... T^m$ (Theorem 3).

It is interesting to observe that no new conditions on w or f are to be added when going from case (b) to case (c).

The conditions are useful in those cases in which the dynamics of the reference string to be generated is not obtained from that of a real string but is designed to have certain features which are deemed desirable or necessary in experimenting with memory policies. The idea of using two or more window sizes in characterizing the dynamics of a program was suggested by the poor accuracy with which artificial strings generated on the basis of a single window size represent the behavior of real strings in non-working-set environments. An interesting theoretical question, for which we have no answer at this moment, is the one regarding the maximum number of window sizes necessary and sufficient to characterize a given reference string fully; this is the minimum value of m such that there is one and only one reference string (the given one) having the corresponding multiple (w, f) characterization.

Acknowledgements

The author is deeply indebted to Masanori Kanazawa for his careful reading of the first draft of the paper, and his pointing out several notational inconsistencies. He is also grateful to Ellen Boyle for her fast and accurate typing of the manuscript.

References

[1] P.J. Denning, "The working set model for program behavior", *Comm. ACM* *11*, 5 (May 1968), 323-333.

[2] P.J. Denning, "Working sets past and present", *IEEE Trans. Software Engineering SE-6*, 1 (January 1980), 64-84.

[3] D. Ferrari, "A generative model of working set dynamics", *Proc. ACM-SIGMETRICS Conf. on Measurement and Modeling of Computer Systems*, Las Vegas, Nevada (Sept. 1981), 52-57.

[4] J.E. Shemer and B. Shippey, "Statistical analysis of paged and segmented computer systems", *IEEE Trans. Electronic Computers EC-15*, 6 (Dec. 1966), 855-863.

[5] N. Faller, "Design and analysis of program behavior models for the reproduction of working-set characteristics", Ph.D. dissertation, University of California, Berkeley; also PROGRES Report No. 81.7, June 1981.

[6] D. Ferrari, "Characterization and reproduction of the referencing dynamics of programs", *PERFORMANCE '81, Proc. 8th Int. Symp. on Computer Performance Modeling, Measurement, and Evaluation*, North-Holland, Amsterdam (Nov. 1981), 363-372.

[7] C.R. Dutt and D. Ferrari, "Characterizing and reproducing program referencing dynamics", PROGRES Report No. 82.4, University of California, Berkeley, April 1982.

[8] T.P. Lee, "Experimental design of a generative model based on working set size characterizations", M.S. Report, University of California, Berkeley; also PROGRES Report No. 82.7, July 1982.

[9] T.P. Lee, "The properties and limiting behavior of working set size strings and flat faults", PROGRES Report No. 82.8, University of California, Berkeley, November 1982.