US005893120A

United States Patent [19]

Nemes

[11] **Patent Number: 5,893,120**

[45] **Date of Patent: Apr. 6, 1999**

[54] **METHODS AND APPARATUS FOR INFORMATION STORAGE AND RETRIEVAL USING A HASHING TECHNIQUE WITH EXTERNAL CHAINING AND ON-THE-FLY REMOVAL OF EXPIRED DATA**

[76] Inventor: **Richard Michael Nemes**, 1432 E. 35th St., Brooklyn, N.Y. 11234-2604

[21] Appl. No.: **775,864**

[22] Filed: **Jan. 2, 1997**

[51] **Int. Cl.**[6] **G06F 17/30**

[52] **U.S. Cl.** **707/206**; 707/1; 707/100; 707/101; 707/202

[58] **Field of Search** 707/1, 200–206, 707/2, 100–103

[56] **References Cited**

U.S. PATENT DOCUMENTS

5,121,495	6/1992	Nemes	707/3
5,202,981	4/1993	Shackelford	707/1
5,287,499	2/1994	Nemes	707/206

OTHER PUBLICATIONS

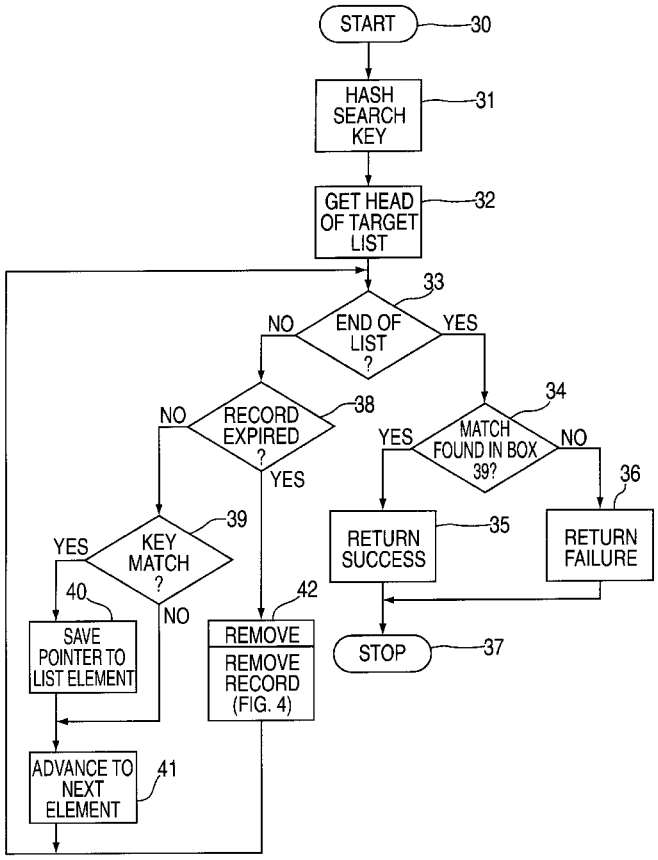
D.E. Knuth, *The Art of Computer Programming,* vol. 3, Sorting and Searching, Addison–Wesley, Reading, Massachusetts, 1973, pp. 506–549.

R.L. Kruse, *Data Structures and Program Design,* Second Edition, Prentice–Hall, Englewood Cliffs, New Jersey, 1987, Section 6.5, "Hashing," and Section 6.6, Analysis of Hashing, pp. 198–215.

D. F. Stubbs and N.W. Webre, *Data Structure with Abstract Data Types and Pascal,* Brooks/Cole Publishing Company, Monterey, California, 1985, Section 7.4, "Hased Implementations," pp. 310–336.

Primary Examiner—Thomas G. Black
Assistant Examiner—Hosain T. Alam

[57] **ABSTRACT**

A method and apparatus for performing storage and retrieval in an information storage system is disclosed that uses the hashing technique with the external chaining method for collision resolution. In order to prevent performance deterioration due to the presence of automatically expiring data items, a garbage collection technique is used that removes all expired records stored in the system in the external chain targeted by a probe into the data storage system. More particularly, each insertion, retrieval, or deletion of a record is an occasion to search an entire linked-list chain of records for expired items and then remove them. Because an expired data item will not remain in the system long term if the system is frequently probed, it is useful for large information storage systems that are heavily used, require the fast access provided by hashing, and cannot be taken off-line for removal of expired data.

8 Claims, 6 Drawing Sheets