STANDARD

11558

First edition
1992-09-01

Information technology — Data compression for information interchange — Adaptive coding with embedded dictionary — DCLZ Algorithm

Technologies de l'information — Compression de données pour l'échange d'information — Codage adaptif avec un dictionnaire incorporé — Algorithme DCLZ

Contents

Foreword

ISO (the International Organization for Standardization) and IEC (the International Electrotechnical Commission) form the specialized system for worldwide standardization. National bodies that are members of ISO or IEC participate in the development of International Standards through technical committees established by the respective organization to deal with particular fields of technical activity. ISO and IEC technical committees collaborate in fields of mutual interest. Other international organizations, governmental and non-governmental, in liaison with ISO and IEC, also take part in the work.

In the field of information technology, ISO and IEC have established a joint technical committee, ISO/IEC JTC 1. Draft International Standards adopted by the joint technical committee are circulated to national bodies for voting. Publication as an International Standard requires approval by at least 75 % of the national bodies casting a vote.

International Standard ISO/IEC 11558 was prepared by the European Computer Manufacturers Association (as Standard ECMA-151) and was adopted, under a special "fast-track procedure", by Joint Technical Committee ISO/IEC JTC 1, *Information technology,* in parallel with its approval by national bodies of ISO and IEC.

Annexes A to C of this International Standard are for information only.

Patents

During the preparation of the ECMA standard, information was gathered on patents upon which application of the standard might depend. Relevant patents were identified as belonging to Hewlett Packard Limited. However, neither ECMA nor ISO/IEC can give authoritative or comprehensive information about evidence, validity or scope of patent and like rights. The patent holders have stated that licences will be granted under reasonable and non-discriminatory terms. Communications on this subject should be addressed to

> Hewlett Packard Limited
> Computer Peripherals Bristol
> Filton road
> Stoke Gifford
> Bristol BS12 6QZ
> United Kingdom

Introduction

In the past decades ISO/IEC have published numerous International Standards for magnetic tapes, magnetic tape cassettes and cartridges, as well as for optical disk cartridges. Those media developed recently have a very high physical recording density. In order to make an optimal use of the resulting data capacity, compression algorithms have been designed which allow a reduction of the number of bits required for the representation of user data in coded form.

In future, these compression algorithms will be registered by an International Registration Authority to be established by ISO/IEC. The registration will consist in allocating to each registered algorithm a numerical identifier. For a recorded medium this should be included in the recorded format to indicate which compression algorithm(s) has been used.

This International Standard is the first of a series of International Standards for compression algorithms.

Adaptive coding with embedded dictionary - DCLZ algorithm

1   Scope

This International Standard specifies a lossless compression algorithm to reduce the number of bits required to represent information coded by means of 8-bit bytes. This algorithm is known as DCLZ (Data Compression according to Lempel and Ziv).

This International Standard specifies neither the strategy for resetting the dictionary nor that for freezing it, as these are implementation-dependent.

This algorithm is particularly useful when information has to be recorded on an inter-changeable medium. Its use is not limited to this application.

2   Conformance

A compression algorithm shall be in conformance with this International Standard if its output data stream satisfies the requirements of clause 7.

3   Normative reference

The following standard contains provisions which, through reference in this text, constitute provisions of this International Standard. At the time of publication, the edition indicated was valid. All standards are subject to revision, and parties to agreements based on this International Standard are encouraged to investigate the possibility of applying the most recent edition of the standard listed below. Members of IEC and ISO maintain registers of currently valid International Standards.

ISO/IEC 11576, *Information technology - Procedure for the registration of algorithms for the lossless compression of data*

4   Definitions

4.1   **Code Value:** An integer in the range 0 to 4 095 that is generated by the compression algorithm.

4.2   **Codeword:** A set of 9, 10, 11 or 12 consecutive bits in the output stream to express a Code Value in binary form.

4.3   **compression ratio:** The number of bits in the input stream of the compression algorithm divided by the number of bits in the output stream of the compression algorithm.

4.4   **dictionary:** A table, comprising 3 832 entries, that is used to retain strings of bytes selected from the input stream. Each entry is identified by a unique Code Value that is greater than 263.

4.5   **empty state:** The state in which no data is in the dictionary.

4.6   **frozen state:** The state in which no further data shall be added to the dictionary.

5   Notations and acronyms

-   Numbers in this International Standard are expressed in decimal notation.

-   EOR: end of record.

6   Algorithm identifier

The numeric identifier of this algorithm in the International Register is 32.