

```
=====  
lzhrecover v0.1 20090526  
=====
```

これは何？

解凍できない lha ファイルから、出来る限りのファイルを救出する作業を補助するツールです。

といっても、解凍できない理由はいろいろあるでしょうから、とても限定的なツールです。あまり期待はできません。

+lhaca もしくは Lhaplus で作られた、途中までしか解凍できない lzh ファイルをある程度復元することができます。
(詳細は以下「後日談」参照)

作った背景

とある 1G 以上ある巨大 lha ファイルを解凍していたところ、半分ぐらいのところエラー発生してしまい、さまざまな解凍ツールを試してみるもやはりエラーになった。「解凍できないファイルを少しでも救出できないか？」ネット上でツール探すも見つからず。幸いバイナリエディタで対応する方法を見つけるが、あまりに巨大な lha ファイルである為、思うようにゆかず、作ったツールです。

自分のケースに使うだけの捨てツールっぽいものなので、わりと手抜きです。他のケースで役に立つかは謎です。

lha のファイルのファイル構造

ものすごく簡略化すると

```
+-----+
| ファイル1 (ヘッダ+データ) | 1レコード
+-----+
| ファイル2 (ヘッダ+データ) |
+-----+
|                               |
|                               |
+-----+
| ファイルn (ヘッダ+データ) |
+-----+
| 0x00(エンドマーク) |
+-----+
```

となっています。
このファイル構造で重要なのは、レコード間で相関関係がまったくないので、単純に、壊れているレコードを特定し、これを削除してしまえば、正常な lha ファイルになる (可能性がある) ということです。

壊れているレコードをどのように特定すればよいのか？

lzrecover(本ツール)では、次の2つのアプローチが行えます。

- A) UNLHA32.DLL で読み込み
最も有名な lha の圧縮 / 解凍ライブラリです。
本ツールでは、このライブラリを使って lzh ファイル中の
ファイルを読み込みます。
エラー検出は厳密ですが一度エラーになると、それ以降の
レコードは解析できません。(オプションで破損ヘッダ読み
飛ばしは可能ですが)
- B) 独自解析
本ツール独自で本当に最低限の解析を行うツールで、
UNLHA32.DLL と比べるとお粗末もしいところですが。
ただし、エラー発生しても次レコードを予測して
無理やり最後まで解析を続けます。
予測といっても、ヘッダ中の圧縮方式 "-lhX-" をさがし
ているだけです。

動作に必要なもの

.NET Framework 2.0

[UNLHA32.DLL で読み込み] を使う場合、以下も必要です。

CalDbase.dll, CalDbase.ini (CalDbase.dll Ver.0.10 に含まれます)
WrapperDLL - CalDbase.NET Website
http://caldbase.sourceforge.jp/old_wiki/index.php?WrapperDLL

UNLHA32.DLL (ulh3265d.exe に含まれます)
Common Archivers Library: UNLHA32.DLL
<http://www.csdinc.co.jp/archiver/lib/unlha32.html>

使い方

1. [lzh 選択] で、壊れた lzh ファイルを選択 (drag&drop でも可)
2. [UNLHA32.DLL で読み込み] で、エラー発生するか確認
一覧に表示されたレコードは、正常と思われる。
(この操作は、本ツールの必須操作ではありませんが、
非力な独自解析を補助する情報が得られます)
3. [独自解析] で、エラー発生するか確認
エラー発生しても最後まで処理継続させる。
このときあきらかにエラーのレコードについては、
チェック OFF になる。
4. 必要であれば、エラーと思われるレコードをチェック OFF する。
5. [lzh 出力] で、新たな lzh を出力する。
6. 出力した lzh を使って、1 から繰り返す。
エラーが無くなれば、おそらく解凍できるはず。
解凍できないようなら、正常に解凍できたファイルの次の

ファイルがエラーレコードなので、これを除外して新たな lzh を作る。
この繰り返し。

解凍するツールは LHMelt がお勧め

LHMelt のダウンロード

<http://www2.nsknet.or.jp/micco/mysoft/lhmelt.htm>

初心者の方や書庫関係に明るくない方は、解らない設定は弄らないほうがいいです。
少なくとも『オプション設定』ダイアログから呼び出せる各 DLL の設定ダイアログ
の項目は弄ってはダメです。LHMelt どころか、他のアプリまで正常に動作しな
くなります。

とあるので注意は必要ですが、UNLHA32.DLL の作者さん作成のツールなので
信頼できると思います。
機能的にも、書庫テストできる・破損ヘッダ読み飛ばしできる・ログが一覧で見られる
と充実しているので、今回のように壊れた lzh ファイルを見るのには最適でした。

他のツールもいくつか試してみたのですが、特殊事情（巨大な lzh ファイルで、かつ
ヘッダ周りがあやしいレコードが含まれている）があっても素直に動いてくれたのは
LHMelt だけでした。

1. [編集 書庫のテスト] で書庫テスト
2. [編集 展開] で解凍

「ヘッダ CRC が存在しません」が発生したとき。
警告であって解凍はできるので無視しても構いません。

「格納ファイルの Huffman コードが壊れています」が発生したとき。
[モード 破損ヘッダを読み飛ばす] をチェック ON していれば
そのファイルをスキップして処理（解凍 / 書庫テスト）を継続してくれます。

後日談

復元できなかったファイル名を見ていて気が付いたのですが、
エラーになっている lzh ファイルは、+Lhaca もしくは Lhaplus で作られた
可能性がありました。（以下以外にも CRC エラーが大量に出ましたし）

『+Lhaca 1.24』について

<http://www2.nsknet.or.jp/micco/notes/soft/Lhaca124.htm>

+Lhaca が作成する書庫については以下の制限が存在します：

- * 作成される書庫は h1 形式ヘッダーです。
- * 基本は上記のとおりなのですが、基本ヘッダーのファイル名項目（のみ）で
表現可能な 255 バイトまで作成を行ってしまうため、自身を含めて取り扱
い不可能な書庫が作成されてしまいます。ヘッダー読み込み時のリカバリーは
行われませんので、当該箇所以降の全てのメンバーは無視されます。

『Lhaplus 1.56』について

<http://www2.nsknet.or.jp/micco/notes/soft/Lhaplus156.htm>
にも同様の指摘があります。

これが原因で読めない lzh ファイルについては、本ツールで
独自解析、lzh 出力 とすれば一部「格納ファイルの Huffman
コードが壊れています」が発生しますが、それ以外は復元できます。

備考

- ・本ツール自身は圧縮 / 解凍機能を持ちません。
- ・元の lzh ファイルは Read オンリーで開いているので、このツールがバグを
もっていたとしても、オリジナルの lzh ファイルには影響ないはず。
- ・唯一 Write しているのは、[lzh 出力] ボタンの処理だけです。
- ・処理速度とかまったくチューニングしてません。
- ・内部 32bit で処理しているので、lzh ファイル 4G までなら扱えるはず。
- ・ヘッダレベルは 1,2 のみ対応しています。

参考にしたもの

LHA の書庫構造

<http://www2m.biglobe.ne.jp/dolphin/lha/lha-header.htm>

開けなくなった L Z H の解凍 : データ復旧のオントラック

http://knowledge.ontrack-japan.com/ontrack_now/20040515_mamechisiki.html

Common Archivers Library: UNLHA32.DLL
の ulh3265d.exe の API.TXT

<http://www.csdinc.co.jp/archiver/lib/unlha32.html>

動作・サポートについて

動作保障・サポート保障はありません。
いかなる責任も負いません。

gounx2 at gmail.com