# 重複っぽいファイルチェックツール

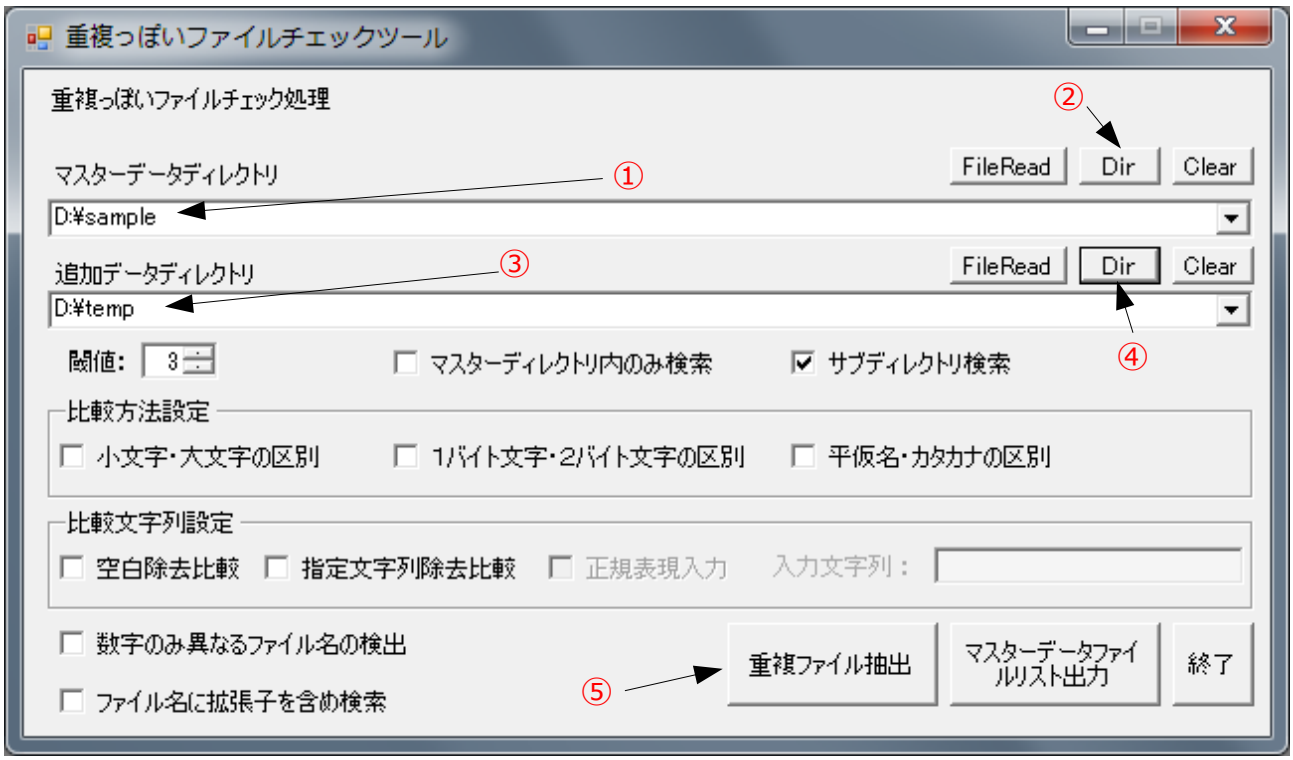
## 操作マニュアル

### 処理概要

当ソフトは、あらかじめ一定の命名規則のもとに名づけられたファイル群を、ディレクトリ単位で比較し重複が疑われるファイルを検出します。また、重複が疑われるファイルを検出した場合、該当ファイルの一覧出力を行うか、ファイルの削除を行うことができます。ファイル削除実行時に実行ディレクトリ下のLogディレクトリに、実行した日付のLOGファイルを作成します。

### 注意事項

当ソフトの処理の一部はファイルの削除を行います。Windowsのシステムフォルダやプログラムフォルダなどのファイルを削除すると、システムに悪影響を与える恐れがあります。データ保存用のフォルダ以外、指定しないようにしてください。当ソフトによって生じた如何なる損害に対しても、作者はその責任を負いません。




## A. 重複ファイル抽出処理実行時（パターンA）

PC内のディレクトリを設定し、該当ディレクトリ内を検索し見つかったファイル群をもとに重複ファイル抽出処理を行います。

1．①マスターディレクトリ設定BOXに比較元のディレクトリを設定

”|”キー(shiftキー＋”¥”)で区切ることによって複数ディレクトリ設定可能

（②の入力フォルダ設定ボタン押下でも設定可能）

2．③追加ディレクトリ設定BOXに比較先のディレクトリを設定

設定に関してはマスターと同様

3．検索・比較オプション設定

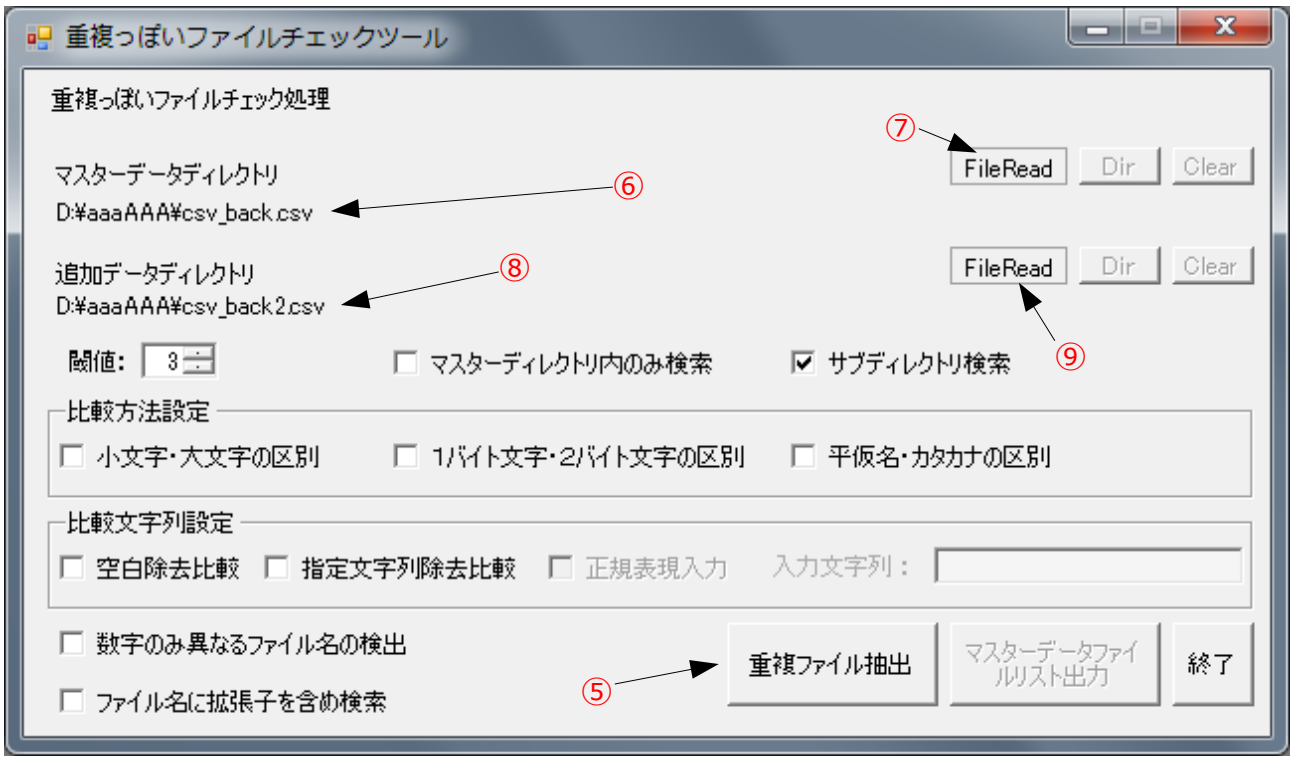
検索・比較オプション内容を設定

※オプション内容はDで後述

4．重複ファイル抽出処理実行

⑤重複ファイル抽出ボタン押下により処理が実行されます。

（E.進捗状況表示画面が表示）




## B. 重複ファイル抽出処理実行時（パターンB）

あらかじめ作成されたファイルリストをもとに重複ファイル抽出処理を行います。ファイルリストがPCの現在状態と一致していると保証できないため、こちらの方法での重複ファイル抽出時は注意が必要です。

※ファイルリスト作成方法はCで後述

1．⑥マスターファイルリスト情報を設定

⑦FileReadボタン押下（入力ファイルを選択）

2．⑧追加ファイルリスト情報を設定

設定に関してはマスターと同様

3．検索・比較オプション設定

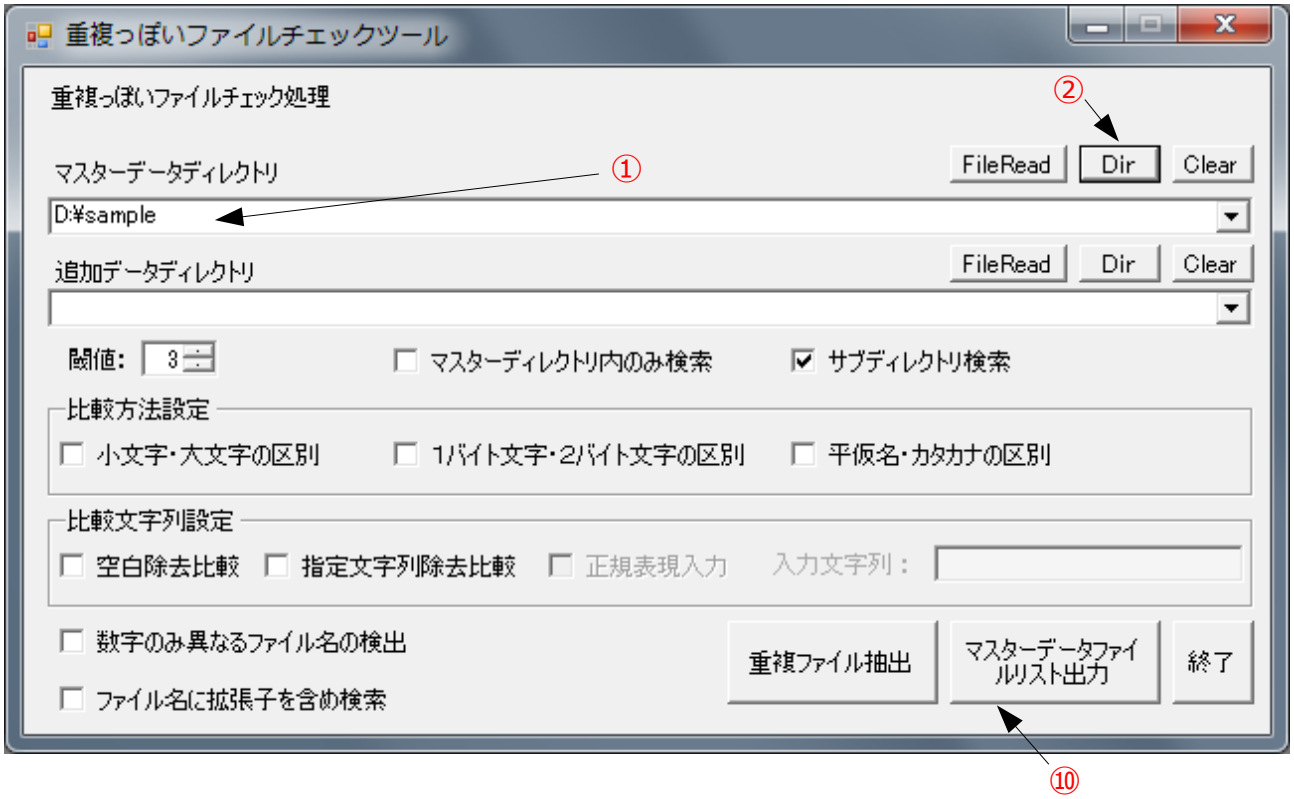
検索・比較オプション内容を設定

※オプション内容はDで後述

4．重複ファイル抽出処理実行

⑤重複ファイル抽出ボタン押下により処理が実行されます。

（E.進捗状況表示画面が表示）




## C. マスターデータファイルリスト出力処理実行時

PC内のディレクトリを設定し、該当ディレクトリ内を検索し見つかったファイル群をもとにファイルリストを作成します。

1．①マスターディレクトリ設定BOXに比較元のディレクトリを設定

”|”キー(shiftキー＋”¥”)で区切ることによって複数ディレクトリ設定可能

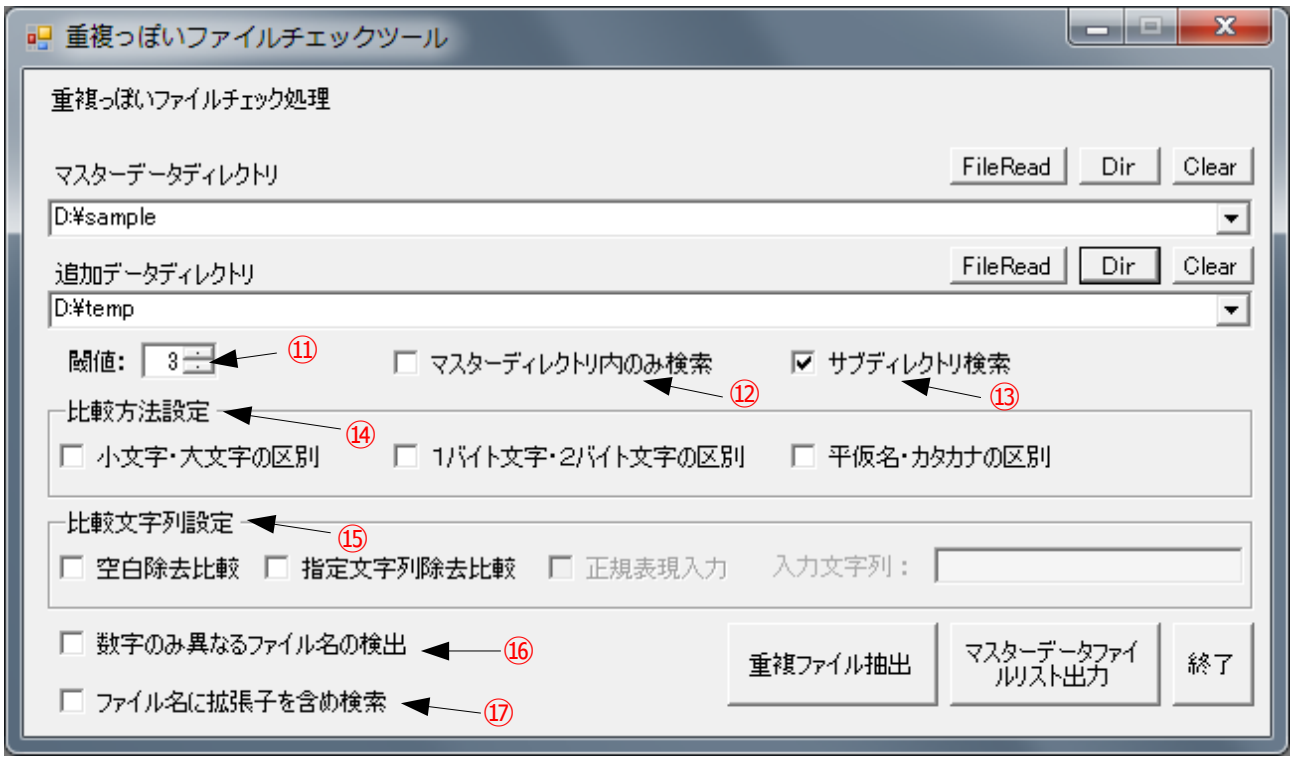
（②の入力フォルダ設定ボタン押下でも設定可能）

2．サブディレクトリ検索設定

サブディレクトリのファイル群をリストに出力するか設定します

3．マスターデータファイルリスト出力処理実行

⑩マスターデータファイルリスト出力ボタン押下により処理が実行されます。




## D．重複ファイル抽出処理実行時オプション設定

1．⑪閾値設定

何文字異なるものまで抽出するかの値を設定

”０”の時は同一のファイル名のみ抽出（”０”の時のみ通常より高速になります）

2．⑫マスターディレクトリ内のみ検索設定

マスターディレクトリ内のみ検索チェックボックスを押下した場合

マスターディレクトリ内でのみファイル比較を行います

追加ディレクトリの設定内容は全てクリアされます

3．⑬サブディレクトリ検索設定

サブディレクトリ検索チェックボックスを押下した場合サブディレクトリのファイル群も検索します

4．⑭比較方法設定

小文字・大文字の区別チェックボックスを押下した場合、小文字・大文字を別の文字として比較します（１バイト文字・２バイト文字の区別，平仮名・カタカナの区別についても同様）

5．⑮比較文字列設定

空白除去比較チェックボックスが押下されている場合、ファイル名からスペースを除いて比較を行います。重複ファイル検出率が上昇し、検出精度が下降しますが、常にチェックすることをおすすめします

指定文字列除去比較チェックボックスがチェックされている場合、テキストボックスに文字列を入力し、その文字列をファイル名から除いた状態で重複確認処理を行います

例）また正規表現入力チェックボックスをチェックしている場合に

¥(.*?¥)　　と入力すると

（）で閉じられた文字列を除去して比較を行います

6．⑯数字のみ異なるファイル名の検出設定

数字のみ異なるファイル名の検出チェックボックスを押下した場合

議事録_100606.txt　，　議事録_100607.txt

このようなファイルも重複ファイルとして検出するようになります（閾値が１以上の時）

※デフォルト（チェックなし）では検出されません

7．⑰ファイル名に拡張子を含め検索設定

ファイル名に拡張子を含め検索チェックボックスを押下した場合

拡張子を含めたファイル名を使用して、重複ファイル抽出処理を行います

## E．進捗状況表示時

重複ファイル抽出処理実行中にこの画面が表示されます。

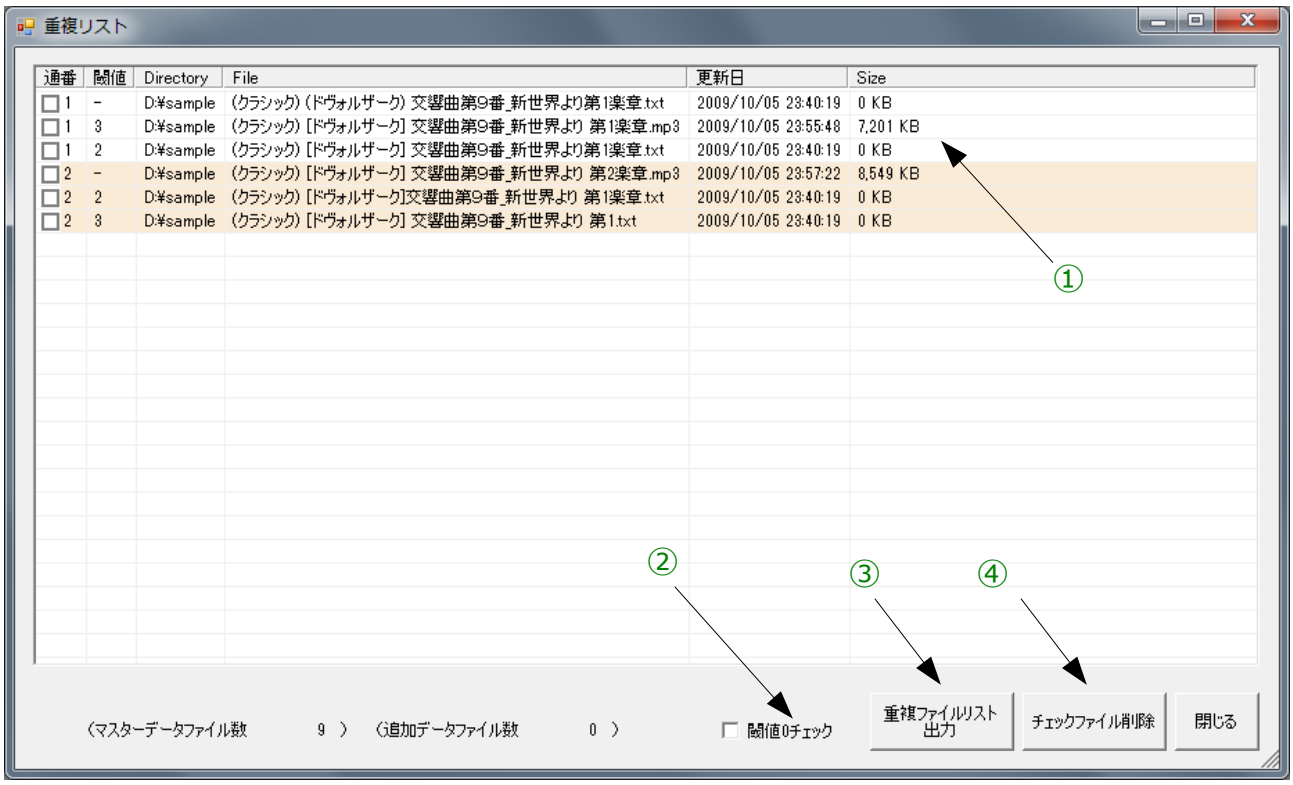
1．重複ファイル抽出処理キャンセル時

①キャンセルボタン押下により処理がキャンセルされます

（②に現在の処理進捗状況が表示されます）

2．重複ファイル抽出処理完了時

自動的にこの画面が消え、重複リスト画面が表示されます




## F．重複ファイルリスト出力時

1．③重複ファイルリスト出力ボタン押下

①重複リストの内容が指定したファイル名でCSV形式で出力されます。

## G．チェックファイル削除時

1．①重複リストの内容確認

（閾値"-"のファイルがマスターのファイルで、同一通番で下に表示されているファイルが、重複可能性ありのファイルです。閾値はファイル名の差異の値）

2．①重複リストから削除ファイル設定

①重複リスト内の削除したいファイルのチェックボックスをチェックします。

（②閾値０チェックをチェックすると閾値０のもの（ファイル名が同一のファイル）が全てチェック状態になります）

（①重複リストの上で右クリックすると[関連付けで開く]，[指定アプリケーションで開く]が選択できます）

3．削除処理実行

④チェックファイル削除ボタン押下により処理が実行されます。

（実行ディレクトリ下のLogディレクトリに、実行した日付のLOGファイルを作成します）