

# TETDM 統合環境 クイックスタートガイド

## ～ツール利用編～

2011.12.16 作成

### 1. 利用する前に

統合環境は、各種テキスト処理ツール、可視化ツールを実行する基盤となるものです。以下の OS での動作を確認しています。

- Windows XP
- Windows Vista
- Windows 7 (32bit, 64bit)
- Mac OS X (Snow Leopard)

統合環境を利用するためには、java の実行環境（JRE: Java Runtime Environment, JDK: Java Development Kit など）、日本語形態素解析器 ChaSen がインストールされている必要があります。これらの入手、インストール方法については下記サイトを参考にしてください。

Java : <http://java.com/ja/>

ChaSen : <http://chasen-legacy.sourceforge.jp/>

#### 1.1. Windows の場合

ChaSen をインストール後、環境変数 PATH を設定する必要があります。コントロール・パネルの [システム] アイテムを起動し、[システムのプロパティ]（Windows7 では [詳細設定]）から [詳細設定タブ] を選択します（図 1）。一番下にある [環境変数] をクリックします（図 2）。

システム環境変数から [Path] を選択し、編集を押して ChaSen へのパスを追加します。

（例）C:¥Program Files¥ChaSen20; ← Path の最後尾に追記

統合環境をダウンロードし、適当なフォルダに Zip ファイルを解凍します。

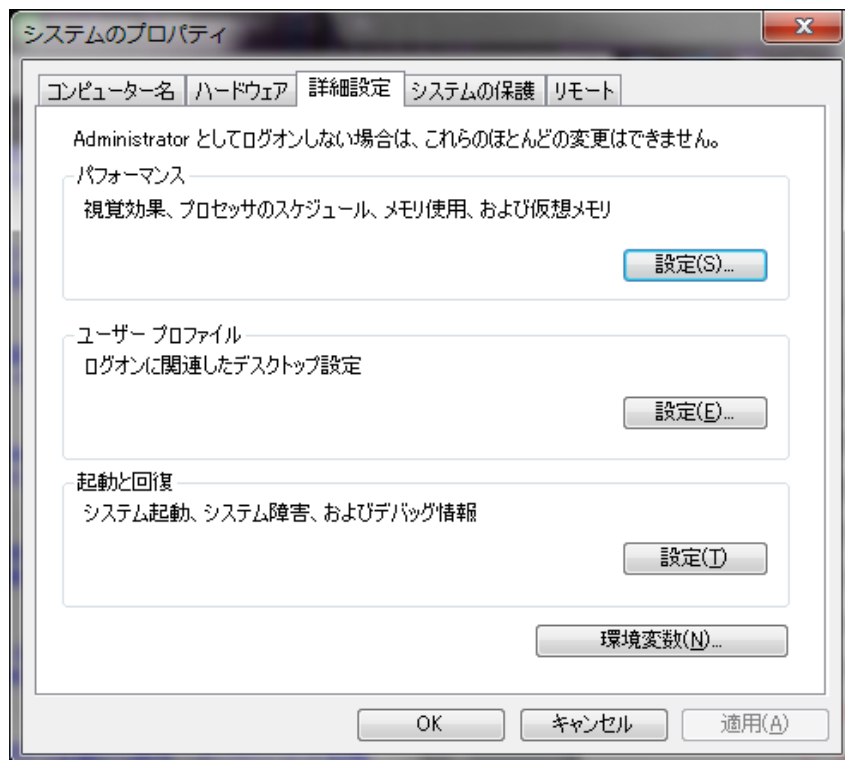


図 1. システムのプロパティ

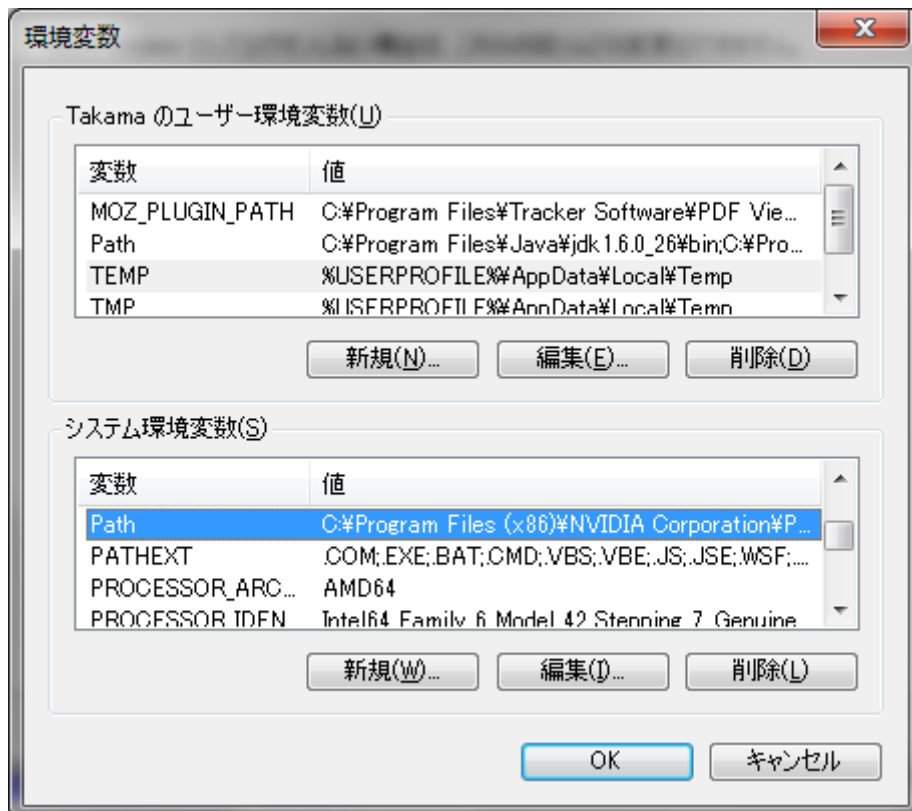


図 2. 環境変数の設定

## 1.2. Mac OS の場合

Macports をインストールし、ターミナルから ChaSen をインストールします。

```
$ sudo port install ChaSen
```

統合環境をダウンロードし、適当なフォルダに Zip ファイルを解凍します。

この状態では ChaSen にパスが通っていないので、統合環境を初めて起動した際に設定を変更する必要があります。手順は「3. 環境設定」を参照してください。

## 2. 実行環境の起動

解凍したフォルダ内にある TETDM.jar をダブルクリックすると、統合環境が起動します（図 1）。最初に必ずしなくてはならないことは、処理対象のテキストファイルを読み込むことです。画面下部、右から 4 番目の「テキスト読込」ボタンを押して読み込むテキストファイルを指定します。デフォルトで、サンプルデータが入っているフォルダ（text）が表示されます（図 2）。名前が X のファイルを選択した場合、統合環境は以下の 2 種類のファイルを生成します。

- X2 … X を複製したファイル
- X2.cha … ChaSen で処理した結果を格納したファイル

※Windows では文字コード Shift JIS, Mac では EUC で作成したテキストファイルを使用してください。

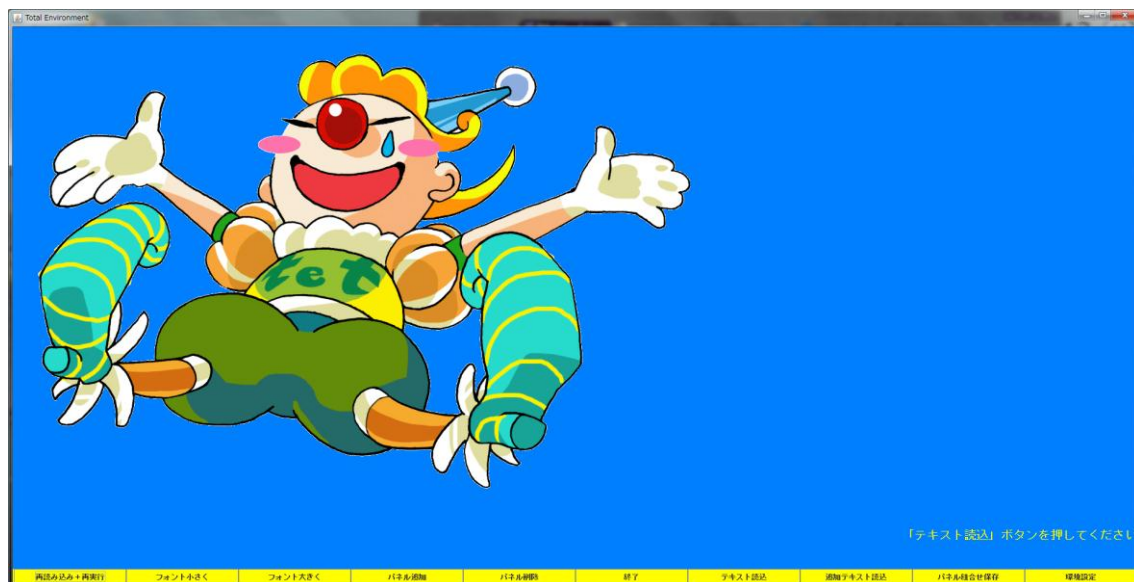


図 1. 統合環境の起動画面

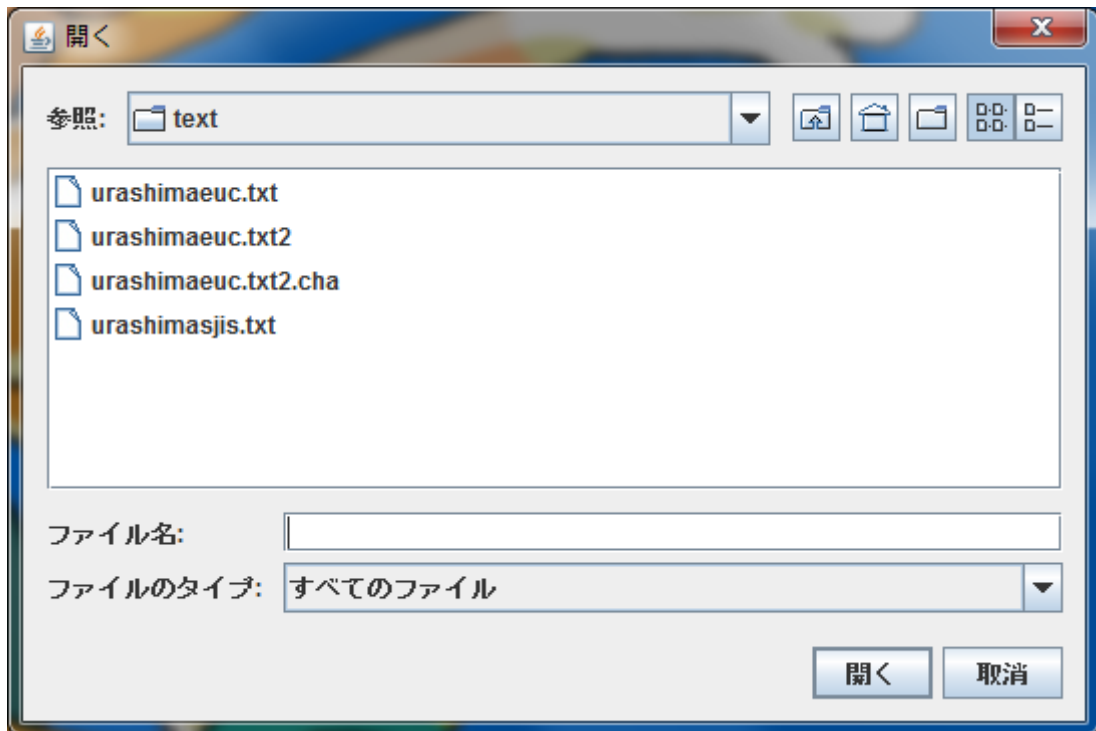


図 2. テキストファイルの選択

図 3 は、Windows で urashimasjis.txt を読み込んだ場合の画面です。デフォルトでは、2 つのパネルが表示されます。どちらも、画面下部にメニューボタンが表示されます。黄色（最下段）は統合環境に対する操作、ピンク（下から 2 段目）は処理ツールの選択、水色（同 3 段目）は可視化ツールの選択に対応したボタンです。下部 2,3 段目はパネル毎に表示されます（図 4）。

処理ツールと可視化ツールには対応関係があり、選択された処理ツールの処理結果を可視化可能なツールのボタンのみが黒字で表示されます。

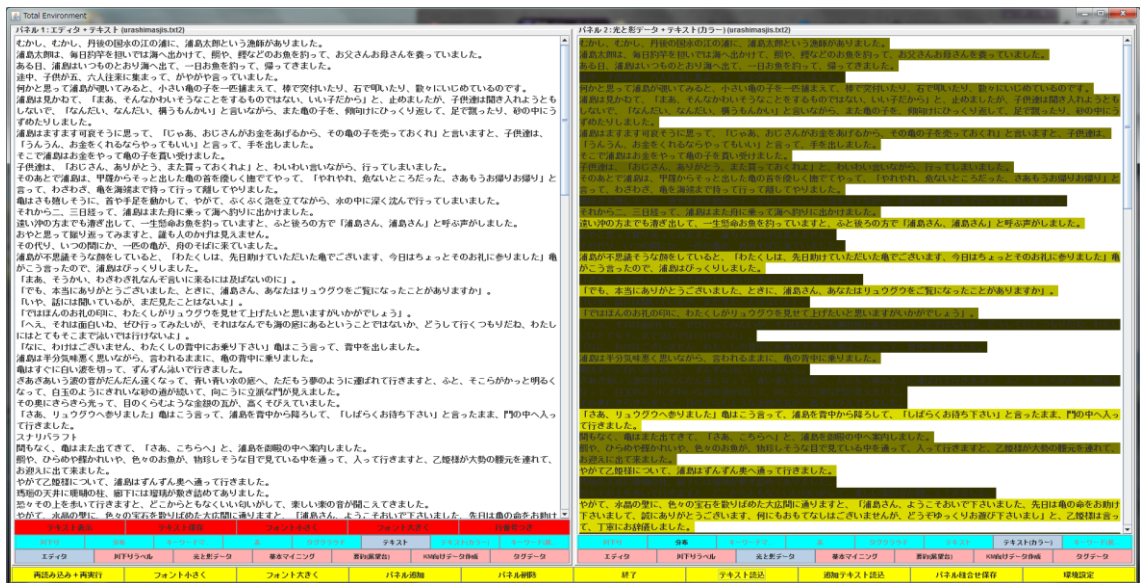


図 3. ファイルを読み込んだところ

メニュー	名前	キーワード	色	タグクラウド	テキスト	テキスト(カラー)	キーワード(色)
エディタ	川下リラベル	光と影データ	基本マイニング	契約(展望台)	KM向けデータ作成		タグデータ

図 4. パネル毎に表示されるメニューボタンの例

統合環境に対する操作（最下部にある黄色のボタン）：

- ・再読み込み + 再実行
  - … 処理対象のテキストファイルを読み込み直します。
- ・フォント小さく / フォント大きく
  - … パネル内に表示されるテキストのフォントサイズを変更します。
- ・パネル追加 … 新しいパネルを右側に追加します。
- ・パネル削除 … 一番右側のパネルを削除します。
- ・終了 … 統合環境を終了します。
- ・テキスト読込 … 処理対象のテキストファイルを指定して読み込みます。指定されたファイルは、表示されている全てのパネルに適用されます。
  - ・追加テキスト読込 … 処理対象のテキストファイルを新たに指定して読み込みます。すでに表示されているパネルには影響を与えず、新たなパネルが生成されて処理結果が表示されます。その後「再読み込み + 再実行」を押すと全てのパネルに指定したテキストファイルが適用されます。
  - ・パネル組み合わせ保存 … 現在表示されているパネルの構成（パネル枚数、各パネルで利用されているツール）を保存します。これにより、次回統合環境を起動してテキスト読込を行うと、同じ構成のパネルが表示されます。

・環境設定 … 統合環境の形態素解析処理や、利用するモジュールの選択を行います。

### 3. 環境設定

環境設定画面（図5）では、形態素解析処理の設定（画面左側）、統合環境にロードする処理ツール、可視化ツールの選択（右側）を行います。設定変更をした場合は「設定保存」ボタンを押してください。「設定初期化」ボタンを押すと、インストール時の初期設定に戻ります。

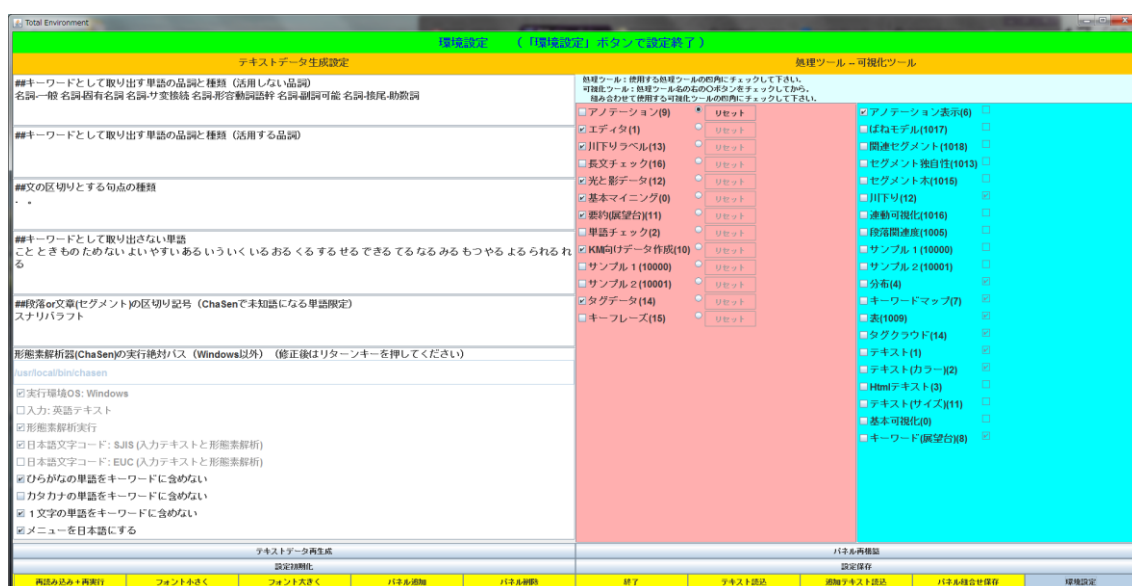


図 5. 環境設定画面

（重要）Mac の場合、MacPorts を利用して ChaSen をインストールした場合には、「形態素解析器（ChaSen）の実行絶対パス（Windows 以外）」に以下を入力し、リターンキーを押した後に設定を保存してください。

**形態素解析器(ChaSen)の実行絶対パス (Windows以外) (修正後はリターンキーを押して**  
**/opt/local/bin/chasen**

図 6. Mac における ChaSen の実行絶対パス指定（必須）