

MACH 1.0을 얻는 방법

1. cs.sungshin.ac.kr/~shim/demo/mach.html에 접속하여 MACH 1.0을 내려 받는다.
2. 내려 받기에 문제가 있는 경우 (가끔 발생한다는 보고가 있음)
cs.sungshin.ac.kr/~shim/demo/download-mach.htm에 접속하여 MACH 1.0을 사용하는 데 따른 조건을 숙지한 후 내려 받기 양식을 채워 이 내용을 첨부하여 shim@sungshin.ac.kr로 메일을 보내 요청할 수도 있다.

MACH 1.0 사용법

참고 : 리눅스(유닉스) 계열의 운영체제를 기준으로 사용법을 설명한다.
윈도즈의 경우도 거의 비슷한 방법으로 사용할 수 있다.

1. 수행 파일

- (a) MACH 1.0에서 이용할 수 있는 옵션과 간단한 설명을 보려면 `-h` 옵션을 사용한다.

```
$ mach -h
```

- (b) 키보드에서 어절(문장)을 읽어들이며 형태소 분석 결과를 화면으로 출력하려면

```
$ mach
MACH 1.0 by Kwangseob Shim
  School of Computer Science and Engineering
  Sungshin Women's University
>
```

- (c) 파일을 읽어 들여 형태소 분석을 하고 결과를 화면으로 출력하려면 분석할 파일명을 준다.

```
$ mach file
$ mach file1 file2
```

(d) 파일을 읽어 들여 형태소 분석을 하고 결과를 파일로 받으려면

```
$ mach file > output
$ mach file1 file2 > output
```

(e) 다중 CPU가 있는 시스템에서 쓰레드를 생성하여 다중 처리를 하려면 `-t` 옵션을 사용한다.

```
$ mach -t4 file1 file2 file3 file4
```

`-t` 뒤에 오는 숫자는 쓰레드의 개수를 나타내는데, 위와 같이 수행할 경우 4개의 쓰레드가 생성되어 각 쓰레드에서 *file1*, *file2*, *file3*, *file4*를 각각 분석하게 된다.

(f) 형태소 분석 결과를 화면으로 출력하지 않도록 하려면 `-s` 옵션을 사용한다. (MACH 1.0의 분석 속도를 측정하는데 사용될 수 있다.)

```
$ mach -s file
$ mach -s file1 file2
```

주의 : 아래 옵션은 구현이 되어 있으나 실험적인 것으로 아직 완전하지 않을 수 있다. (`-j`, `-e` 옵션의 경우 아직 사전 정비가 끝나지 않아 부분적으로만 가능하며, 나중에 시간이 나면 사전 정비를 할 예정임)

(g) MACH 1.0은 복합 명사를 단위 명사로 분석하는 것을 기본으로 하는데 복합 명사를 단위 명사로 분석하지 않기를 원한다면 `-n` 옵션을 사용한다. (단, 이미 복합명사가 사전에 등재된 경우에는 단위 명사로 부

```
$ mach
한국은행에서 → [한국 + 은행 + 에서]로 분석
```

```
$ mach -n
한국은행에서 → [한국은행 + 에서]로 분석
```

(h) MACH 1.0은 동사/형용사화 접미사를 별도의 형태소로 분리하는 것을 기본으로 하는데 접미사를 선행하는 체언과 결합하여 하나의 용언으로 분석되도록 하려면 `-v` 옵션을 사용한다.

\$ mach

공부하다 → [공부 + 하 + 다]로 분석

\$ mach -v

공부하다 → [공부하 + 다]로 분석

- (i) MACH 1.0은 복합 조사를 단위 조사로 분리하는 것을 기본으로 하는데 복합 조사를 단위 조사로 분리되지 않도록 하려면 **-j** 옵션을 사용한다.

\$ mach

학교에서는 → [학교 + 에서 + 는]으로 분석

\$ mach -j

학교에서는 → [학교 + 에서는]으로 분석

- (j) MACH 1.0은 어미에 보조사가 결합된 경우 이를 분리하는 것을 기본으로 하는데 이를 원하지 않는 경우 **-e** 옵션을 사용한다.

\$ mach

먹는데도 → [먹 + 는데도]로 분석

\$ mach -e

먹는데도 → [먹 + 는데 + 도]로 분석

2. 라이브러리

연구용 라이브러리 공개시 추가할 예정.