

# Sendic Search Engine 0.2 API

## C# Version (only SENDIC 2.0 interface )

본 문서는 SSE(Sendic Search Engine) Project 를 진행 중인 R&P 에서 작성 하였으며 개방형 검색엔진인 SSE 0.2을 이용하여 자신에게 맞는 응용 프로그램을 작성 할 수 있는 API 에 대해서 설명 하고 있습니다.

SSE는 계속해서 개발이 진행 되고 있기 때문에 본 문서는 꾸준히 변경 될 수 있습니다.

최신 문서는 <http://kldp.net/projects/sendic> 에서 확인 할 수 있습니다.

본 문서는 2009.12.5일에 작성된 문서입니다.

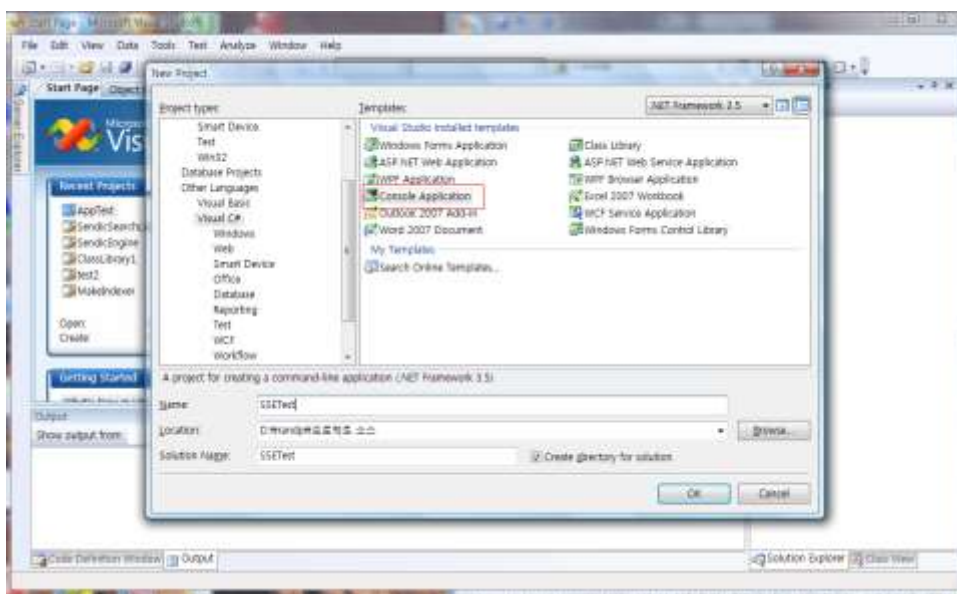
-R&P 

## 1. 개발 환경

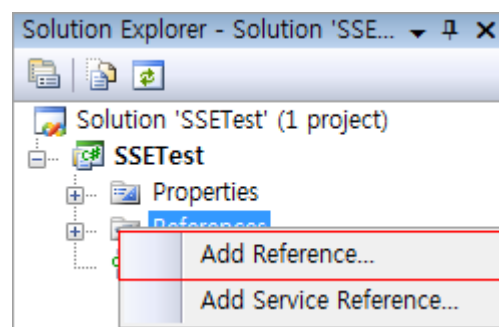
제공된 c# dll 형태나 기타 c# 라이브러리 형태를 사용 할 수 있는 개발 환경에서 적용 할 수 있다.

여기서는 Microsoft visual studio 2008의 c# 관련 프로젝트를 기준으로 설명 하겠다.

- (1) visual studio 2008을 실행 시켜서 SSE 라이브러리를 적용시킬 기존의 프로젝트를 열거나 새로운 프로젝트를 실행 한다. 여기서는 Console Application Project 를 생성한다.



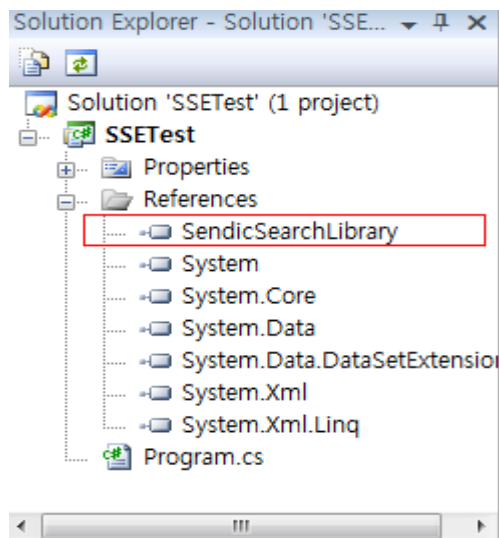
- (2) 생성된 프로젝트에서 제공된 라이브러리를 참조 할 수 있도록 추가 한다. 여기서는 제공된 dll 라이브러리를 추가 하는 것을 기준으로 설명 한다. visual studio 의 solution explorer 에 Reference 부분을 볼 수 있다. Reference 부분을 선택한 후 마우스 오른쪽 버튼을 클릭 하고 Add Reference 를 클릭한다.



나타나는 창에서 Browse 를 선택 하고 제공된 SendicSearchLibrary.dll 를 선택 하게 되면 SSE 라이브러리가 프로젝트에 추가 된다.



추가된 라이브러리는 solution explorer 에서 확인 할 수 있다.



## 2. Sendic Class (using SendicEngine.App)

### (1) 개요

Sendic Class 는 SSE API를 이용하여 SENDIC 2.0 프로그램을 제작하기 위하여 구현된 class 이다. 따라서 Sendic Class 를 분석하면 SSE API에 대해서 어느 정도 파악할 수 있다.

### (2) 멤버 메소드

Define	explanation
Sendic(string filename)	생성자. filename= 검색할 파일명.
void Search(string str)	검색을 한다. str= 검색할 단어
void GetPage(int page)	Search에서 검색한 결과를 page 단위로 접근 할 수 있도록. 가리키는 역할을 한다. page = 가리켜야 하는 page (1부터)
void GetPageTranslate(int page)	GetPage와 비슷한 역할을 하며 Semantic 데이터도 구하게 된다. page = 가리켜야 하는 page (1부터)
string[] FindCntSents(int wcnt)	정해진 개수의 word 로 이루어진 sentence를 100개까지 구한다. wcnt = sentence 를 구성하는 word 의 개수 return = 검색한 sentence의 string 배열
void SelBlockData(int indicate)	block 에서 특정 데이터를 가리키다. indicate = 가리키는 데이터의 순서 번호 (0부터)
void MovePreBlock()	검색해서 지금 가리키고 있는 데이터에서 이전 블록으로 jump 하여 이전 블록 데이터를 접근할 수 있게 한다.
void MoveNextBlock()	검색해서 지금 가리키고 있는 데이터에서 다음 블록으로 jump 하여 다음 블록 데이터를 접근할 수 있게 한다.
void MoveBegin()	Search 한 초기 블록으로 다시 접근 할 수 있도록 초기 블록을 가리킨다.
SDATA GetSDATA()	검색한 데이터에 접근 할 수 있는 SDATA class 를 구한다. return = 검색한 데이터에 접근할 수 있는 SDATA
void FindSem(string str)	의미 데이터를 검색,

	sequence iterator(Pre(), Next())과 같이 사용 str= 검색할 문장이나 단어
bool Pre()	sequence iterator 의 이전 데이터를 구한다. return = true: 이전 검색 데이터가 있을 경우. false: 이전 검색 데이터가 없을 경우.
bool Next()	sequence iterator 의 다음 데이터를 구한다. return = true: 다음 검색 데이터가 있을 경우. false: 다음 검색 데이터가 없을 경우.
void MyAdd(string filename, string str)	지정된 파일에 개인 데이터를 저장. filename = 개인 사용자 데이터 저장 하는 파일명 str = 저장할 개인데이터
void MyAdd(string filename, string str)	지정된 파일에 개인 데이터를 삭제. filename = 삭제할 개인 사용자 데이터를 저장하고 있 는 파일명 str = 삭제할 개인데이터
void MySearch(string filename, string query)	지정된 개인 데이터 파일에서 검색을 한다. GetMyData()와 같이 사용된다. filename = 검색할 개인사용자 파일 query = 검색할 문자열
String[] GetMyData()	MySearch 에서 검색한 결과를 문자열로 구한다. 연속으로 호출 하면 계속 다음 데이터를 구한다. return : 검색한 문자열 , null= 검색결과가 없을 경우
String[] GetMyShow (string filename)	지정된 파일명의 개인 사용자 데이터를 구한다. 연속으로 호출하면 계속해서 다음 데이터를 구한다. return: 구한 개인사용자 문자열 null= 결과가 없을 경우

(3) 멤버 변수

name	explanation
_count	검색한 데이터의 수
_TotalPage	검색한 데이터에 대한 총 페이지 수

### 3. SDATA Class (using SendicEngine.App)

(1) 개요

검색된 결과 데이터를 접근 할 수 있는 iterator 역할을 한다.

(2) 멤버 메소드

define	explanation
SDATA(int size)	생성자. size = 저장 관리할 sentence 의 개수
int GetLen()	검색한 데이터 중 현재 가리키고 있는 page 에서의 sentence 수를 구한다. return = 검색된 sentence 수
string[] GetSentens()	검색한 sentence 를 구함 return = 검색한 sentence 의 string 배열
string[] GetTranslation()	검색한 sentence에 대한 해석 데이터를 구함 return = 검색한 sentence 에 대한 해석데이터 string 배열
string[] GetPreBlock()	현재 가리키고 있는 검색된 데이터의 이전 블록을 구함 return= 검색된 이전 블록의 sentence string 배열
string[] GetNextBlock()	현재 가리키고 있는 검색된 데이터의 다음 블록을 구함 return= 검색된 다음 블록의 sentence string 배열

(3) 멤버 변수

\_len = 검색된 sentence 수

\_PreLen = 이전 블록 sentence 수

\_NextLen = 다음 블록 sentence 수

\_tlen = 해석 데이터 수

## 4. Sample Code

### (1) 검색된 데이터 구하기

검색 기능을 구현하기 위해서는 SSE에서 지원하는 index 구조로 구성된 index 파일이 있어야 한다. filename.index, filename.sen , filename.kor(선택사항)  
Make indexing 기능은 Sendic Class 에서 지원하지 않고 별도로 제공된다.

### Script.sdc 검색

```
namespace SSETest
{
    class Program
    {
        static void Main(string[] args)
        {
            /******
            * searching hair.
            * *****/

            SDATA sdata; // define SDATA.
            Sendic sendic = new Sendic("Script.sdc"); // create Sendic Class instance , using filename
            string[] str; // save sentences.
            sendic.Search("hair"); // searching "hair"
            sendic.GetPage(1); // indicate 1 page.
            sdata = sendic.GetSDATA(); // get data iterator
            str = sdata.GetSentens(); // get sentence in 1 page

            // display sentence
            for (int i = 0; i < sdata._len; i++)
            {
                Console.WriteLine(" sentence " + i + " : " + str[i]);
            }
        }
    }
}
```



```
C:\Windows\System32\cmd.exe
sentence 0 : rachel gets up and ruffles ross's hair and then walks off. ross now
sentence 1 : in her wild hair and hot pants.
sentence 2 : fingernails are painted dark red. her hair is puffed out
sentence 3 : queen isabel of spain. the light in her blonde hair
sentence 4 : water. as he turns, we see his face. disfigured by a hair
sentence 5 : the ship's boy is singing to himself. despite his hair
sentence 6 : alonso glances towards the ship's boy, with his hair lip.
sentence 7 : standing at the prow of the ship, his long dark hair
sentence 8 : against a tree. he has removed his helmet. his hair
sentence 9 : with a moldering tuft of human hair attached to the
sentence 10 : hair spills out of the helmet.
sentence 11 : sits tom welles, middleaged, hair neat, suit crisp and
sentence 12 : black hair and grey eyes...
sentence 13 : san crawls on his belly across the metal sheeting, his hair blown
in the
sentence 14 : the father manages a smile for everybody. he ruffles the little b
oy's hair
sentence 15 : john smith, the hair on his domed forehead singing off, regards t
hen.
sentence 16 : he is young, has dark curly hair and rugged outlines to his
sentence 17 : ahead in a state of near catatonia. hair tangled and wild. but at
least she's not so blue
sentence 18 : blows some hair out of her eyes.
sentence 19 : without warning he grabs her by the hair and hurls her against th
```

## (2) get pre block of first sentence

검색한 sentence 중에 특정한 sentence 를 선택해서 pre block 을 구할 수 있다.

Script.sdc 검색 데이터 중 첫 번째 sentence 의 pre block 구함

```

/*****
 * get sentence in pre block.
 * *****/
SDATA sdata; // define SDATA.
Sendic sendic = new Sendic("Script.sdc"); // create Sendic Class instance , using filename.
string[] str; // save sentences.
sendic.Search("hair"); // searching "hair"
sendic.GetPage(1); // indicate 1 page.
sdata = sendic.GetSDATA(); // get data iterator
sendic.SetBlockData(0); // indicate first data.
str = sdata.GetPreBlock();
// display sentence
for (int i = 0; i < sdata._PreLen; i++)
{
    Console.WriteLine(" sentence " + i + " : " + str[i]);
}

```

C:\Windows\System32\cmd.exe

```

sentence 0 : rachel gets up and ruffles ross hair and then walks off. ross nu...
sentence 1 : in her wild hair and hot pants.
sentence 2 : fingernails are painted dark red. her hair is puffed out
sentence 3 : queen isabel of spain. the light in her blonde hair
sentence 4 : water. as he turns, we see his face, disfigured by a hair
sentence 5 : the ships boy is singing to himself, despite his hair.
sentence 6 : allows glance towards the ships bay, with his hair lip.
sentence 7 : standing at the prow of the ship, his long dark hair
sentence 8 : against a tree. he has removed his helmet. his hair
sentence 9 : with a moldering tuft of human hair attached to the
sentence 10 : hair spills out of the helmet.
sentence 11 : sits tom wellen, middleaged. hair neat, suit crisp and
sentence 12 : black hair and grey eyes...
sentence 13 : can crawl on his belly across the metal sheeting, his hair blow...
in the
sentence 14 : the father manages a smile for everybody. he ruffles the little b...
eyes hair
sentence 15 : john smith, the hair on his doned forehead singing off. regards t...
men.
sentence 16 : he is young, has dark curly hair and rugged outlines to his
sentence 17 : ahead in a state of near catatonia. hair tangled and wild. but at...
least shes not so blue
sentence 18 : blows some hair out of her eyes.
sentence 19 : without warning he grabs her by the hair and huris her against th...

```

C:\Windows\System32\cmd.exe

```

sentence 0 : ross probably.
sentence 1 : rachel really?
sentence 2 : ross but, but, you know ill tell you something, percuss is way m...
sentence 3 : what?
sentence 4 : rachel yeah, right!
sentence 5 : ross it is. or eventually it kinda of hoves out, but hopefully wh...
sentence 6 : you are left with is trust and security and or, well in the com...
sentence 7 : of my ex-wife. . . . .backhauled no you know far all those
ross 8 : people who miss out on that so, gazian thing, theres all that
ross 9 : cher good stuff.
rachel sighs sh.
ross but no, i dont think thats gonna be you. . . .
rachel you dont ?
ross um, no i see, or. big. . . . . gudies in your future.
rachel really?
ross um
ross 10 : rachel we do?
ross 11 : ross i do.
sentence 12 : rachel oh ross. your so great. ross.
sentence 13 : rachel gets up and ruffles ross hair and then walks off. ross no...

```



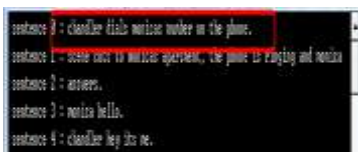
### (3) jump block

검색한 sentence 중에 특정한 sentence 를 선택해서 pre block 이나 next block 로 Jump 하여 연속적으로 pre block 이나 next block 를 구할 수 있다.

Script.sdc 검색 데이터 중 첫 번째 sentence 의 pre block jump 하여서 연속적으로 pre block 를 구함

```
/******  
 * jump block  
 * *****/  
  
SDATA sdata; // define SDATA.  
Sendic sendic = new Sendic("Script.sdc"); // create Sendic Class instance , using filename.  
string[] str; // save sentences.  
sendic.Search("hair"); // searching "hair"  
sendic.GetPage(1); // indicate 1 page.  
sdata = sendic.GetSDATA(); // get data iterator  
sendic.SetBlockData(0); // indicate first data.  
str = sdata.GetPreBlock();  
sendic.MovePreBlock(); // first jump pre block  
sendic.MovePreBlock(); // second jump pre block  
sendic.SetBlockData(0);  
// display sentence  
str = sdata.GetPreBlock();  
for (int i = 0; i < sdata.PreLen; i++)  
{  
    Console.WriteLine(" sentence " + i + " : " + str[i]);  
}
```

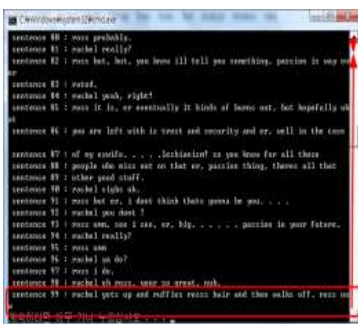
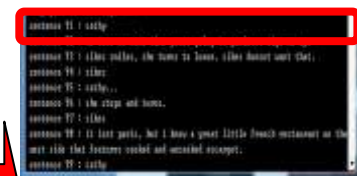
Pre Block



First jump pre block



Second jump pre block



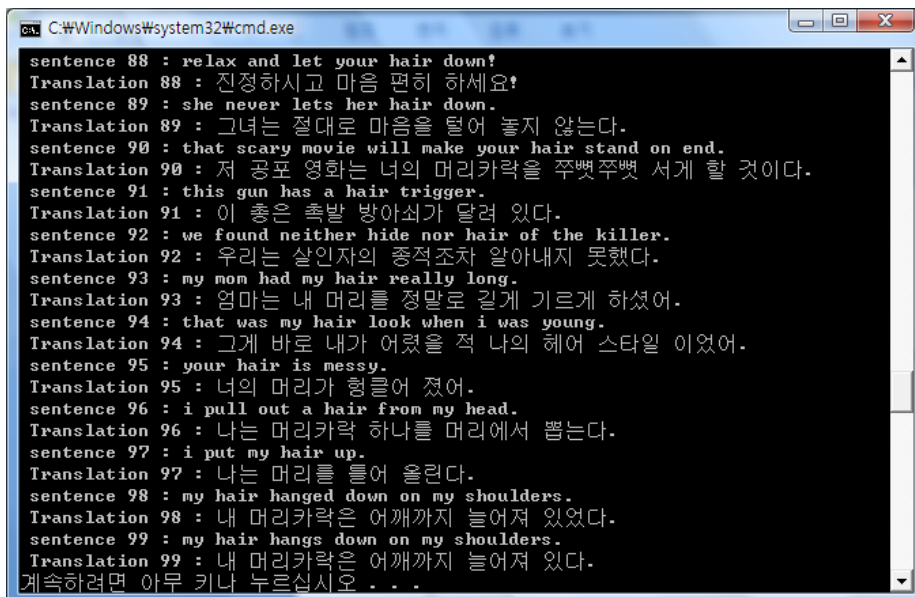
#### (4) 해석 데이터 구하기

검색한 sentence 와 관계 있는 해석 데이터를 구하여 사용 할 수 있다.

Trans.sdc 검색 하여 sentence 와 해석 데이터를 같이 구함

```
/*
 *
 * translation data
 *
 * *****/

SDATA sdata; // define SDATA.
Sendic sendic = new Sendic("Trans.sdc"); // create Sendic Class instance , using filename.
string[] str, str2; // save sentences.
sendic.Search("hair"); // searching "hair"
sendic.GetPageTranslate(1); // indicate 1 page. using Translation data.
sdata = sendic.GetSDATA(); // get data interator
str = sdata.GetSentens(); // get sentence
str2 = sdata.GetTranslation(); // get translation data.
for (int i = 0; i < sdata._len; i++)
{
    Console.WriteLine(" |sentence " + i + " : " + str[i]);
    Console.WriteLine(" Translation " + i + " : " + str2[i]);
}
}
```



```
C:\Windows\system32\cmd.exe

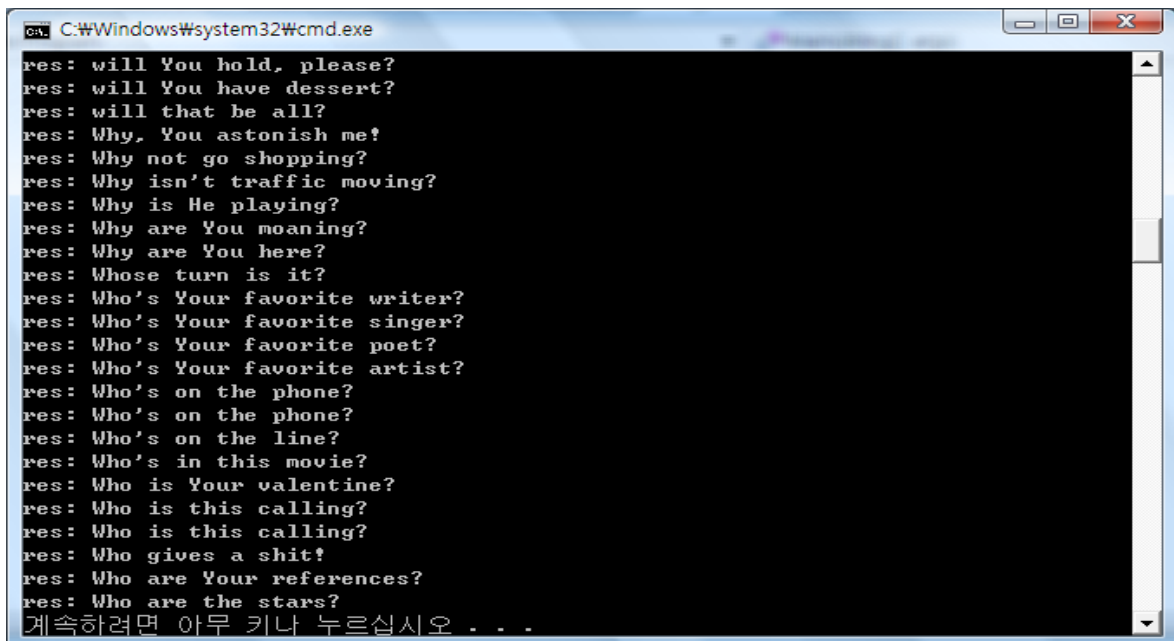
sentence 88 : relax and let your hair down!
Translation 88 : 진정하시고 마음 편히 하세요!
sentence 89 : she never lets her hair down.
Translation 89 : 그녀는 절대로 마음을 털어 놓지 않는다.
sentence 90 : that scary movie will make your hair stand on end.
Translation 90 : 저 공포 영화는 너의 머리카락을 쭈뼛쭈뼛 서게 할 것이다.
sentence 91 : this gun has a hair trigger.
Translation 91 : 이 총은 촉발 방아쇠가 달려 있다.
sentence 92 : we found neither hide nor hair of the killer.
Translation 92 : 우리는 살인자의 종적조차 알아내지 못했다.
sentence 93 : my mom had my hair really long.
Translation 93 : 엄마는 내 머리를 정말로 길게 기르게 하셨어.
sentence 94 : that was my hair look when i was young.
Translation 94 : 그게 바로 내가 어렸을 적 나의 헤어 스타일 이었어.
sentence 95 : your hair is messy.
Translation 95 : 너의 머리가 헝클어 졌어.
sentence 96 : i pull out a hair from my head.
Translation 96 : 나는 머리카락 하나를 머리에서 뽑는다.
sentence 97 : i put my hair up.
Translation 97 : 나는 머리를 틀어 올린다.
sentence 98 : my hair hanged down on my shoulders.
Translation 98 : 내 머리카락은 어깨까지 늘어져 있었다.
sentence 99 : my hair hangs down on my shoulders.
Translation 99 : 내 머리카락은 어깨까지 늘어져 있다.
계속하려면 아무 키나 누르십시오 . . .
```

#### (5) 지정된 개수의 단어 문장 구하기

지정된 수의 단어로 이루어진 문장을 구할 수 있다.

Trans.sdc 에서 4개의 단어로 이루어진 문장을 구함

```
Sendic sendic = new Sendic("Trans.sdc");  
  
string[] str=sendic.FindCntSents(4); // 4개의 단어로 설정  
  
if (str != null) // 반드시 결과가 있는지 체크 해봐야 함  
{  
    for (int i = 0; i < str.Length; i++)  
    {  
        Console.WriteLine("res: " + str[i]);  
    }  
}  
sendic.Close();
```



```
ca. C:\Windows\system32\cmd.exe  
res: will You hold, please?  
res: will You have dessert?  
res: will that be all?  
res: Why, You astonish me!  
res: Why not go shopping?  
res: Why isn't traffic moving?  
res: Why is He playing?  
res: Why are You moaning?  
res: Why are You here?  
res: Whose turn is it?  
res: Who's Your favorite writer?  
res: Who's Your favorite singer?  
res: Who's Your favorite poet?  
res: Who's Your favorite artist?  
res: Who's on the phone?  
res: Who's on the phone?  
res: Who's on the line?  
res: Who's in this movie?  
res: Who is Your valentine?  
res: Who is this calling?  
res: Who is this calling?  
res: Who gives a shit!  
res: Who are Your references?  
res: Who are the stars?  
계속하려면 아무 키나 누르십시오 . . .
```

## (6) 개인 사용자 데이터 저장, 검색

지정된 파일에 데이터를 저장하고 검색할 수 있다. (색인을 하지 않음)

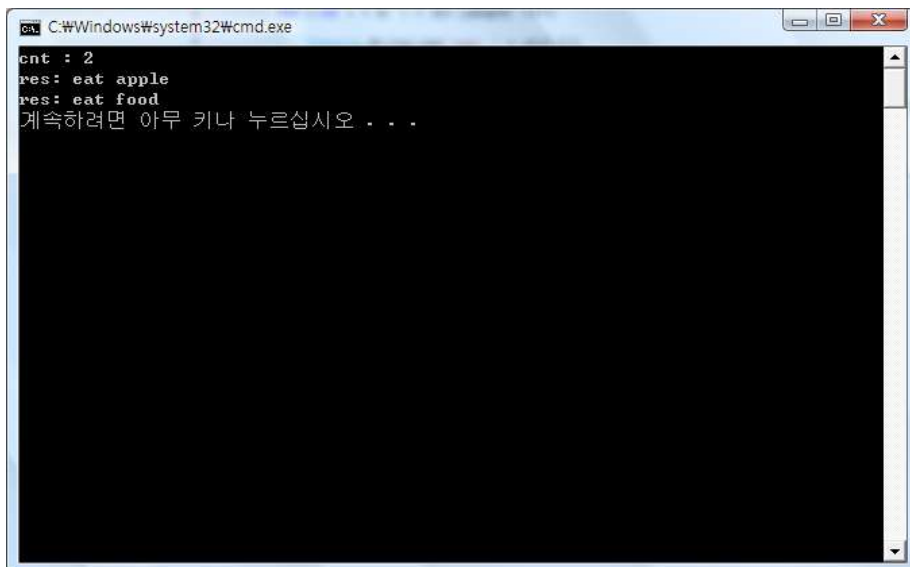
agun.txt 에 데이터를 저장하고 eat 로 검색함

```
Sendic sendic = new Sendic();
sendic.MyAdd("agun.txt", "eat apple"); // add eat apple in agun.txt
sendic.MyAdd("agun.txt", "two line"); // add two line in agun.txt
sendic.MyAdd("agun.txt", "eat food"); // add eat food in agun.txt

sendic.MySearch("agun.txt", "eat"); // Search eat from agun.txt
string[] str = sendic.GetMyData(); // Get search data.

if (str != null)
{
    Console.WriteLine("cnt : " + str.Length);
    for (int i = 0; i < str.Length; i++)
    {
        Console.WriteLine("res: " + str[i]);
    }
}

// 여기서 str = sendic.GetMyData(); 를 한번 더 호출 하면 검색 결과중 다음 결과를 가져온다.
sendic.Close(); // 데이터 초기화
```



```
C:\Windows\system32\cmd.exe
cnt : 2
res: eat apple
res: eat food
계속하려면 아무 키나 누르십시오 . . .
```

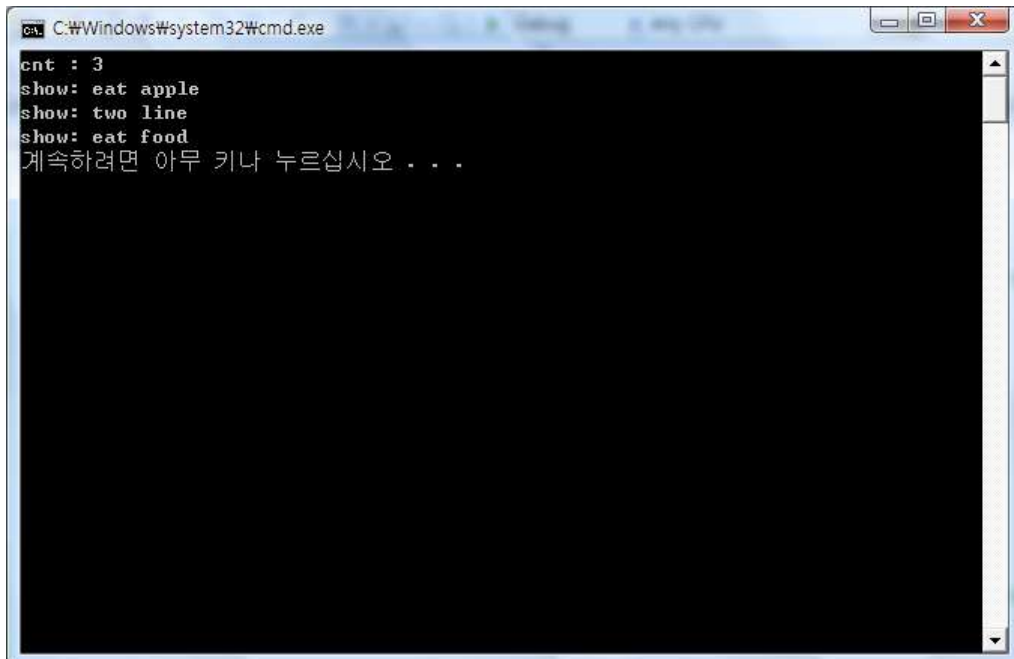
## (7) 개인 사용자 데이터 보기

저장된 데이터 보기.

agun.txt 에 데이터를 저장되어 있는 데이터를 봄

```
Sendic sendic = new Sendic();
string[] str = sendic.GetMyShow("agun.txt");
if (str != null)
{
    Console.WriteLine("cnt : " + str.Length);
    for (int i = 0; i < str.Length; i++)
    {
        Console.WriteLine("show: " + str[i]);
    }
}

//여기서 str = sendic.GetMyShow(); 를 한번 더 호출 하면 검색 결과중 다음 결과를 가져온다.
sendic.Close();
```



```
C:\Windows\system32\cmd.exe
cnt : 3
show: eat apple
show: two line
show: eat food
계속하려면 아무 키나 누르십시오 . . .
```

## (8) 의미(해석) 데이터 검색

한글로 해석데이터 중에 검색 한다. 현재 단일검색만 지원

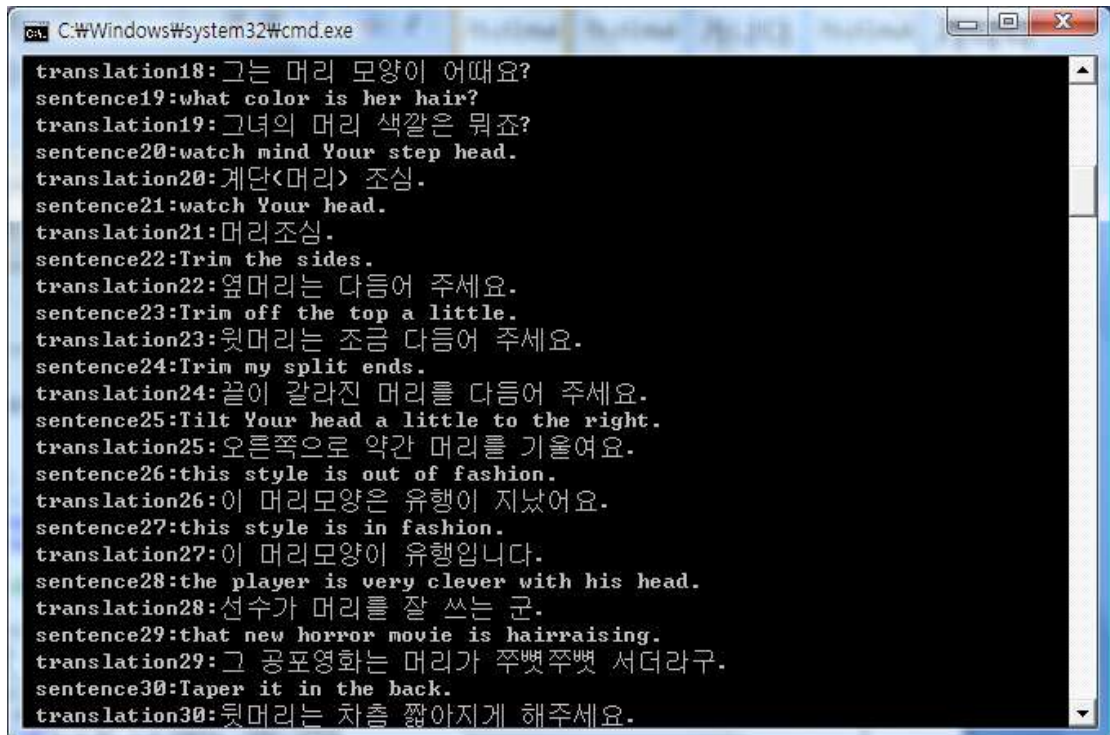
Trans.sdc 에서 "머리" 해석데이터를 검색함

```
SDATA sdata;
Sentic sendic = new Sentic("Trans.sdc");
string[] str;
string[] str2;

// 사용자에게 입력받은 값이 한글인지 체크 하고 아래와 같이 실행 하면 된다.

sendic.FindSem("머리");
sdata = sendic.GetSDATA();

str = sdata.GetSentens();
str2 = sdata.GetTranslation();
for (int i = 0; i < sdata._len; i++)
{
    Console.WriteLine(" sentence" + i + ":" + str[i]);
    Console.WriteLine(" translation" + i + ":" + str2[i]);
}
sendic.Close();
```



```
translacion18:그는 머리 모양이 어때요?
sentence19:what color is her hair?
translacion19:그녀의 머리 색깔은 뭐죠?
sentence20:watch mind Your step head.
translacion20:계단<머리> 조심.
sentence21:watch Your head.
translacion21:머리조심.
sentence22:Trim the sides.
translacion22:옆머리는 다듬어 주세요.
sentence23:Trim off the top a little.
translacion23:윗머리는 조금 다듬어 주세요.
sentence24:Trim my split ends.
translacion24:끝이 갈라진 머리를 다듬어 주세요.
sentence25:Tilt Your head a little to the right.
translacion25:오른쪽으로 약간 머리를 기울여요.
sentence26:this style is out of fashion.
translacion26:이 머리모양은 유행이 지났어요.
sentence27:this style is in fashion.
translacion27:이 머리모양이 유행입니다.
sentence28:the player is very clever with his head.
translacion28:선수가 머리를 잘 쓰는 군.
sentence29:that new horror movie is hairraising.
translacion29:그 공포영화는 머리가 쭈뼛쭈뼛 서더라구.
sentence30:Taper it in the back.
translacion30:윗머리는 차츰 짧아지게 해주세요.
```

### (9) 해석 데이터에서 sequence iterator 이용하기

위 (8)번의 다음 데이터들을 Next() 를 사용하여 구할 수 있다. (Pre() 이전)

Trans.sdc 에서 "머리" 해석데이터를 검색하고 Next()를 사용하여 다음 100개의 데이터를 구함

```
SDATA sdata;
Sendic sendic = new Sendic("Trans.sdc");
string[] str;
string[] str2;

// 사용자에게 입력받은 값이 한글인지 체크 하고 아래와 같이 실행 하면 된다.

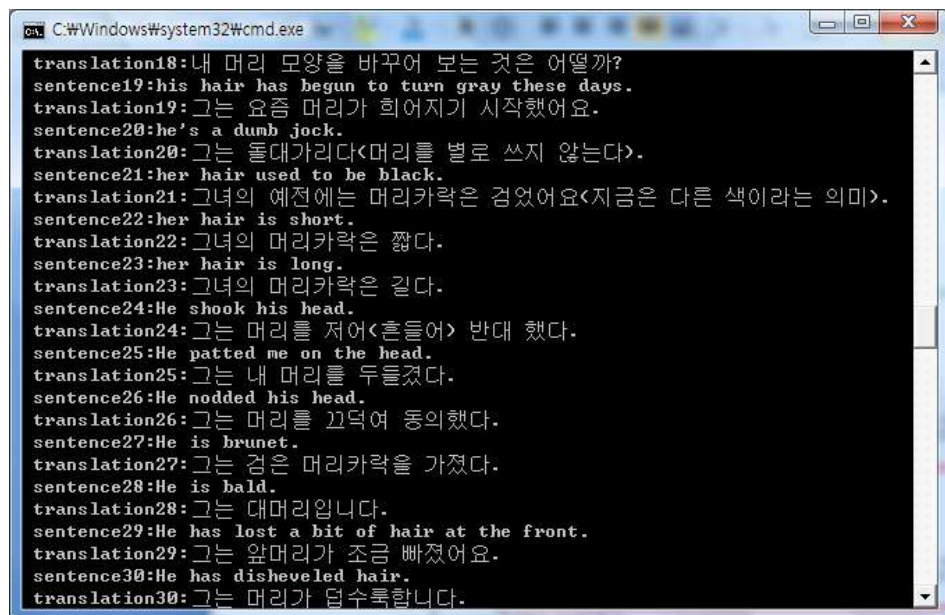
sendic.FindSem("머리");
sdata = sendic.GetSDATA();

str = sdata.GetSentens();
str2 = sdata.GetTranslation();
for (int i = 0; i < sdata._len; i++)
{
    Console.WriteLine(" sentence" + i + ":" + str[i]);
    Console.WriteLine(" translation" + i + ":" + str2[i]);
}

sendic.Next(); // 다음 데이터 구하기 (100개)

str = sdata.GetSentens();
str2 = sdata.GetTranslation();
for (int i = 0; i < sdata._len; i++)
{
    Console.WriteLine(" sentence" + i + ":" + str[i]);
    Console.WriteLine(" translation" + i + ":" + str2[i]);
}

sendic.Close();
```



```
ca. C:\Windows\system32\cmd.exe
translation18:내 머리 모양을 바꾸어 보는 것은 어떨까?
sentence19:his hair has begun to turn gray these days.
translation19:그는 요즘 머리가 희어지기 시작했어요.
sentence20:he's a dumb jock.
translation20:그는 돌대가리다<머리를 별로 쓰지 않는다>.
sentence21:her hair used to be black.
translation21:그녀의 예전에는 머리카락은 검었어요<지금은 다른 색이라는 의미>.
sentence22:her hair is short.
translation22:그녀의 머리카락은 짧다.
sentence23:her hair is long.
translation23:그녀의 머리카락은 길다.
sentence24:He shook his head.
translation24:그는 머리를 저어<흔들어> 반대 했다.
sentence25:He patted me on the head.
translation25:그는 내 머리를 두들겼다.
sentence26:He nodded his head.
translation26:그는 머리를 끄덕여 동의했다.
sentence27:He is brunet.
translation27:그는 검은 머리카락을 가졌다.
sentence28:He is bald.
translation28:그는 대머리입니다.
sentence29:He has lost a bit of hair at the front.
translation29:그는 앞머리가 조금 빠졌어요.
sentence30:He has disheveled hair.
translation30:그는 머리가 덩수룩합니다.
```