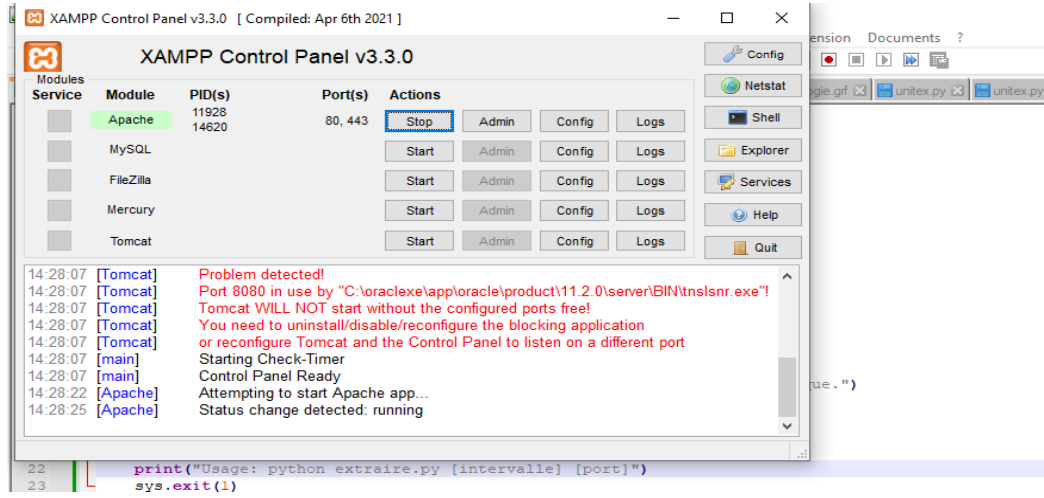


# PROJECT EXTRA INFO

## Etape 1



Ce PC > WIN\_764pr (C:) > Utilisateurs > USER > AppData > Local > Unitex-GramLab > App >

	Nom	Modifié le	Type	Taille
	assembly	09/12/2023 00:05	Dossier de fichiers	
	disclaimers	09/12/2023 00:05	Dossier de fichiers	
	lib	09/12/2023 00:05	Dossier de fichiers	
	licenses	09/12/2023 00:05	Dossier de fichiers	
	manual	09/12/2023 00:06	Dossier de fichiers	
	Alphabet.txt	06/08/2021 01:12	Document texte	1 Ko
	corpus-medical.txt	08/12/2023 16:26	Document texte	1 749 Ko
	Dela_fr.bin	06/08/2021 01:12	VLC media file (.bi...	4 435 Ko
	Dela_fr.inf	06/08/2021 01:12	Informations de c...	305 Ko
36)	enrichir.py	26/12/2023 17:25	Python File	4 Ko
	extraire.py	26/12/2023 11:50	Python File	2 Ko
	Gramlab.jar	06/08/2021 01:14	Executable Jar File	706 Ko
	Norm.txt	06/08/2021 01:12	Document texte	1 Ko
	pom.xml	06/08/2021 01:14	Document XML	2 Ko
	posologie.grf	27/12/2023 17:58	Fichier GRF	2 Ko
	sqlite.py	25/12/2023 21:51	Python File	2 Ko
	Unitex.ico	06/08/2021 01:14	Icône	4 Ko
	Unitex.jar	06/08/2021 01:14	Executable Jar File	1 466 Ko
	unitex.py	27/12/2023 15:26	Python File	1 Ko
	UnitexToolLogger.exe	06/08/2021 06:05	Application	2 966 Ko
	XAlign.jar	06/08/2021 01:14	Executable Jar File	1 142 Ko

```
C:\Users\USER\AppData\Local\Unitex-GramLab\App>extraire.py
Usage: python extraire.py [intervalle] [port]

C:\Users\USER\AppData\Local\Unitex-GramLab\App>_
```

```
C:\Users\USER\AppData\Local\Unitex-GramLab\App>extraire.py Z-A 80
Les lettres de l'intervalle doivent etre en ordre alphabetique.

C:\Users\USER\AppData\Local\Unitex-GramLab\App>_
```

```
C:\Users\USER\AppData\Local\Unitex-GramLab\App>extraire.py a-z 80

Format d'intervalle incorrect:

-Utillisez le format: B-H, E-S, A-W, ou A-Z...ect:

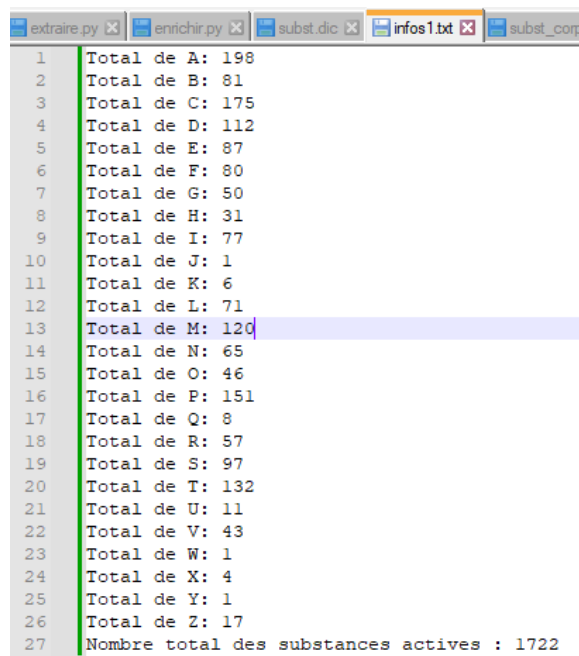
-Lettres en ordre alphabetique , ex A-Z pas Z-A:

-Lettres en majuscules:

C:\Users\USER\AppData\Local\Unitex-GramLab\App>
```

```
C:\Users\USER\AppData\Local\Unitex-GramLab\App>extraire.py A-Z 8081
Extraction réussie. Les résultats sont sauvegardés dans subst.dic et infos1.txt.

C:\Users\USER\AppData\Local\Unitex-GramLab\App>_
```



Letter	Total
A	198
B	81
C	175
D	112
E	87
F	80
G	50
H	31
I	77
J	1
K	6
L	71
M	120
N	65
O	46
P	151
Q	8
R	57
S	97
T	132
U	11
V	43
W	1
X	4
Y	1
Z	17
Total des substances actives	1722

```
1691 vinflunine,.N+subst
1692 vinorelbine,.N+subst
1693 vinpocétine,.N+subst
1694 virus amaril,.N+subst
1695 virus de la varicelle,.N+subst
1696 vismodégib,.N+subst
1697 voriconazole,.N+subst
1698 vorinostat,.N+subst
1699 vortioxétine,.N+subst
1700 warfarine,.N+subst
1701 ximélagatran,.N+subst
1702 xipamide,.N+subst
1703 xylène,.N+subst
1704 xylométazoline,.N+subst
1705 yohimbine,.N+subst
1706 zafirlukast,.N+subst
1707 zalcitabine,.N+subst
1708 zaléplone,.N+subst
1709 zanamivir,.N+subst
1710 ziconotide,.N+subst
1711 zidovudine,.N+subst
1712 zinc,.N+subst
1713 zinc oxyde,.N+subst
1714 zinc peroxyde,.N+subst
1715 ziprasidone,.N+subst
1716 zofénopril,.N+subst
1717 zolmitriptan,.N+subst
1718 zolpidem,.N+subst
1719 zonisamide,.N+subst
1720 zopiclone,.N+subst
1721 zotépine,.N+subst
1722 zuclopenthixol,.N+subst
1723
```

---

## Etape 2

```
C:\Users\USER\AppData\Local\Unitex-GramLab\App>enrichir.py
Usage: python enrichir.py 'corpus-medical.txt'

C:\Users\USER\AppData\Local\Unitex-GramLab\App>enrichir.py jj
Error: Input file must be 'corpus-medical.txt'

C:\Users\USER\AppData\Local\Unitex-GramLab\App>enrichir.py corpus-medical.txt
Enrichissement réussi. Les résultats sont sauvegardés dans subst.dic, subst_corpus.dic, infos2.txt, et infos3.txt.
```

---

extraire.py x enrichir.py x subst.dic x infos1.txt x subst\_corpus.dic x infos2.txt x infos3.txt x

1910 époétine thêta, .N+subst  
1911 époétine zêta, .N+subst  
1912 éprazinone, .N+subst  
1913 éprosartan, .N+subst  
1914 éribuline, .N+subst  
1915 érythromycine, .N+subst  
1916 ésoméprazole, .N+subst  
1917 ésérine, .N+subst  
1918 étamsylate, .N+subst  
1919 étanercept, .N+subst  
1920 ételcalcétide, .N+subst  
1921 éthacridine, .N+subst  
1922 éthambutol, .N+subst  
1923 éthanol, .N+subst  
1924 éthinylestradiol, .N+subst  
1925 éthionamide, .N+subst  
1926 éthosuximide, .N+subst  
1927 éthylmorphine, .N+subst  
1928 étifoxine, .N+subst  
1929 étiléfrine, .N+subst  
1930 étodolac, .N+subst  
1931 étofibrate, .N+subst  
1932 étomidate, .N+subst  
1933 étonogestrel, .N+subst  
1934 étoposide, .N+subst  
1935 étoricoxib, .N+subst  
1936 étravirine, .N+subst  
1937 évolocumab, .N+subst  
1938 évérolimus, .N+subst  
1939 ézétimibe, .N+subst  
1940 plavix, .N+subst  
1941

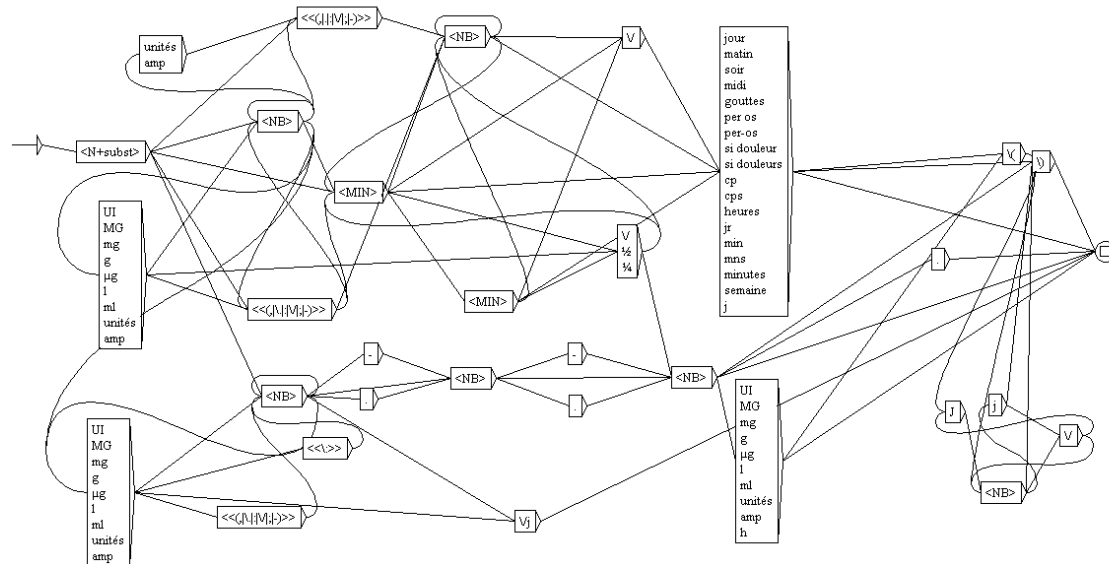
extraire.py x enrichir.py x subst.dic x infos1.txt x subst\_corpus.dic x infos2.txt x infos3.txt x

742 laroxyl, .N+subst  
743 stilnox, .N+subst  
744 tenormine, .N+subst  
745 zanidip, .N+subst  
746 temesta, .N+subst  
747 allopurinol, .N+subst  
748 piascledine, .N+subst  
749 tenormine, .N+subst  
750 lercan, .N+subst  
751 allopurinol, .N+subst  
752 temesta, .N+subst  
753 seropram, .N+subst  
754 furosémide, .N+subst  
755 seropram, .N+subst  
756 inxium, .N+subst  
757 heptamyl, .N+subst  
758 innohep, .N+subst  
759 proctolog, .N+subst  
760 fortimel, .N+subst  
761 transipeg, .N+subst  
762

```
extraire.py x enrichir.py x subst.dic x infos1.txt x subst_corpus.dic x infos2.txt x infos3.txt x
1 Nombre de médicaments dans A: 21
2 Nombre de médicaments dans B: 6
3 Nombre de médicaments dans C: 19
4 Nombre de médicaments dans D: 11
5 Nombre de médicaments dans E: 13
6 Nombre de médicaments dans F: 13
7 Nombre de médicaments dans G: 10
8 Nombre de médicaments dans H: 7
9 Nombre de médicaments dans I: 8
10 Nombre de médicaments dans J: 2
11 Nombre de médicaments dans K: 6
12 Nombre de médicaments dans L: 14
13 Nombre de médicaments dans M: 13
14 Nombre de médicaments dans N: 9
15 Nombre de médicaments dans O: 5
16 Nombre de médicaments dans P: 16
17 Nombre de médicaments dans Q: 0
18 Nombre de médicaments dans R: 6
19 Nombre de médicaments dans S: 20
20 Nombre de médicaments dans T: 22
21 Nombre de médicaments dans U: 1
22 Nombre de médicaments dans V: 3
23 Nombre de médicaments dans W: 0
24 Nombre de médicaments dans X: 4
25 Nombre de médicaments dans Y: 0
26 Nombre de médicaments dans Z: 7
27 Nombre total de médicaments issus de corpus medical: 236
28
```

```
extraire.py x enrichir.py x subst.dic x infos1.txt x subst_corpus.dic x infos2.txt x infos3.txt x
1 Nombre de médicaments enrichis dans A: 17
2 Nombre de médicaments enrichis dans B: 5
3 Nombre de médicaments enrichis dans C: 18
4 Nombre de médicaments enrichis dans D: 11
5 Nombre de médicaments enrichis dans E: 13
6 Nombre de médicaments enrichis dans F: 10
7 Nombre de médicaments enrichis dans G: 10
8 Nombre de médicaments enrichis dans H: 6
9 Nombre de médicaments enrichis dans I: 8
10 Nombre de médicaments enrichis dans J: 2
11 Nombre de médicaments enrichis dans K: 6
12 Nombre de médicaments enrichis dans L: 14
13 Nombre de médicaments enrichis dans M: 12
14 Nombre de médicaments enrichis dans N: 8
15 Nombre de médicaments enrichis dans O: 4
16 Nombre de médicaments enrichis dans P: 15
17 Nombre de médicaments enrichis dans Q: 0
18 Nombre de médicaments enrichis dans R: 5
19 Nombre de médicaments enrichis dans S: 18
20 Nombre de médicaments enrichis dans T: 22
21 Nombre de médicaments enrichis dans U: 1
22 Nombre de médicaments enrichis dans V: 2
23 Nombre de médicaments enrichis dans W: 0
24 Nombre de médicaments enrichis dans X: 4
25 Nombre de médicaments enrichis dans Y: 0
26 Nombre de médicaments enrichis dans Z: 7
27 Nombre total de médicaments issus de l'enrichissement: 218
28
```

# Etape 3



Text DELA FSGraph Lexicon-Grammar XAlign File Edition Windows Help Info

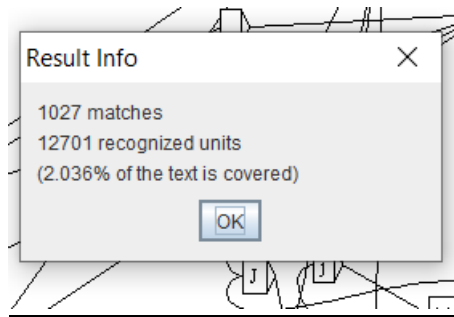
**Lexical Resources**

Select the dictionaries to be applied. You can sort them one by one using the arrows. Note that system dictionaries are given to the Dico program before the user ones.

User resources	System resources
subst.bin	CR.fst2
	Dela_fr.bin
	Dnum-ch-fst-text.fst2
	Dnum-ch.fst2
	Dnum-ord.fst2
	Dnum-tronq.fst2
	Dnum.fst2
	Elements.fst2
	Extrait-DelquefM2.bin
	motsGramf-.bin
	NPr+.fst2
	...

Right-click a dictionary to get information about it :

Clear Default Set Default Apply



Unitex-GramLab IDE 3.3 (rev. 7) (August 09, 2007) - current language is French

Text DELA FSGraph Lexicon-Grammar XAlign File Edition Windows Help Info

Concordance: C:\Users\slak001\AppData\Local\Unitex-GramLab\App\corpus-medical\_snft\concord.html

1010 matches

ml si déglutition OK - Atarax 50 mg/j - [Ulcars 1 s x 3/j](#). \* Sous-cutané : - Prévention de la mal  
 J10) \* Per os : - Lodoz 2.5 mg 1 cp/j - [Ulcars 1 sachet x 2/j](#) - Diffu K 2 cps x 3/j - Coversyl 2  
 nothérapie - Calciparine 0,2 ml x 2/j - [Ulcars 2/j](#) - Amlor 5 1-0-0 - Haldol 10-0-10 gouttes Prob  
 ic : 75 mg/j - Topalgic : 50 mg x 3/j - [Ulcars : 1 sachet x 4/j](#) - Lasilix : 40 mg x 3/j - Vitami  
 eures, créatinine à 6.6 mmol/24 heures, [urée à 198 mmol/24 heures](#), protéinurie à 551 mg/24 heur  
 ement suivant : SUBUTEX 1 matin et soir [VALIUM 10 1 le soir](#) 2 au coucher ATARAX 25 1 matin et m  
 50 : 1 le matin INIPOMP 20 : 1 le soir [VALIUM 10 : 1 le matin](#), 1 le midi, 1 le soir, 1 au couc  
 e suite à la suppression de la prise de [VALIUM le matin](#) et le midi. Au total, patiente de 43 an  
 lline 4 g x 2 par jour IVL 60 minutes - [Vancomycine 1 g par jour](#) IVSE - Gentamycine 240 mg, réa  
 - TRIFLUCAN 400 mg, 2 fois par jour, - [VANCOMYCINE 1 g, 3 fois par jour](#), - PERFALGAN 1 g, 4 fo  
 50 mg/j (J10/J21) - Meronem 500 x 2/j - [Vancomycine 1.5 mg/24 h \(J10/J10\)](#) (IVSE) \* Sonde nasoga  
 al sera retiré. Une bi-antibiothérapie ([Vancomycine 2 g par jour](#) et Rifampycine 1,8 g par jour)  
 g w 3/jour (J5), CIFLOX 20 mg x 2/jour, [VANCOMYCINE 2 g/24h \(J5\)](#), Aérosols BRICANYL ATROVENT x  
 le Docteur GUILLAUME qui la traite par [VANCOMYCINE pendant 24 heures](#) puis RIFAMPICINE. Découve  
 le Docteur GUILLAUME qui la traite par [VANCOMYCINE pendant 24 heures](#) puis RIFAMPICINE. Découve  
 rofuryl 200 1.0.0 Contramal LP100 0.0.1 [Vastarel 35 0.0.1](#) Neurontin 400 0.0.3 En vous remercian  
 ture col femoral droit. TAD : Ceres 2/j [vastarel 35 2/j kaleorid 600 1/j](#) moduretic 1/2 /j dolip  
 jour - Kardégic 75 mg : 1 sachet/jour - [Vastarel 35 mg : 1 cp x 2/jour](#) - Prozac 1 mg : 1 cp/jou  
 ui Traitement: Tenordate 1/j Plavix 1/j [Vastarel 35: 2/j](#) Constantes Commentaires: POTISEK Thoma  
 t de sortie Iopidine : 1 goutte le soir [Xalatan : 1 goutte le soir](#) Contramal 50 : 1 le matin de  
 EXIUM 20 mg/ jour ZOLPIDIEM : 1 le soir [XANAX 0.25 mg le soir](#) RAMIPRIL 10 mg : 1/ jour COKENZEN  
 rologique : arrêt du PERMIXON, mis sous [XATRAL LP 10 le soir](#). Sortie avec sonde urinaire. Consu  
 /2009 17:54 edouard (Médecin) A valider [xyzall 5mg 1 cp le matin](#) discotrine 1 patch le matin 10  
 tin et soir smecta 1 matin midi et soir [xyzall 5mg 1 cp le matin](#) discotrine 1 patch le matin Ap  
 s le soir, EQUANIL 400 : 3 fois / jour, [XYZALL, 1 cp le soir](#). Nous vous prions de croire, Cher  
 ) XXX TENORMINE 1001/jour STABLON 1 0 1 [zanidip 20 1 0 0](#) TEMESTA 2.5 2.5cp le soir ALLOPURINOL  
 metformine 1000: 4/jr topalgic 50: 3/jr [zanidip 20: 1/jr](#) bisoprolol 10: 1/jr simvastatine: 1/jr  
 i douleurs, SKENAN 30 mg matin et soir, [ZELITREX 500 mg : 1/jour](#), ATARAX 25 mg : 1 le soir. En  
 , reconstruire en fin de semaine le T0 - [Zelitrex 500 mg x 2 par jour](#) - Atarax 25 mg ½ au couche  
 réanimation pour lequel elle reçoit du [Zelitrex pendant 10 jours 500 mg x 2 par jour](#). Il n'y a  
 ré remplis x2/j par voie sous cutanée - [ZESTRIL 20 mg : 1 cp le matin](#) - TAHOR 10 mg : 1 cp/j -  
 le matin - TAHOR 10 mg : 1 cp le soir - [ZESTRIL 20 mg : 2 cps le matin](#) - CERIS 20 mg : 1 cp le  
 : lasilix 40 1-0-0 levothyrox 150 1-0-0 [zoloft 50 1-0-0](#) allopurinol 300 1-0-0 diffu k 1-0-0 glu  
 ements. GLUCOPHAGE 850 : 3 cps par jour [ZOLOFT 50 mg : 1 cp](#) à 8 h 00, 1 cp à 20 h 00 INIPOMP 20  
 , un le midi Ø INIPOMP 20, un le soir Ø [ZOLOFT 50, un matin](#) et soir Ø LOVENOX 0,4, une injectio  
 jour le matin, AMAREL 10 1 cp le matin, [ZOLPIDIEM 10 mg 1 cp](#) au coucher, NORMACOL un lavement pa  
 g : 1 matin et soir INEXIUM 20 mg/ jour [ZOLPIDIEM : 1 le soir](#) XANAX 0.25 mg le soir RAMIPRIL 10  
 e : Cardensiel 1.25 : ½ matin et soir - [Zyloric 100 : 1 le midi](#) - Inexium 20 : 1 le midi - Iské  
 .25 : 1 le matin Inexium 20 : 1 le soir [Zyloric 100 : 1 le midi](#) Perfusions d'albumine En restan  
 matin - Aprovel 150 : 1 matin et soir - [Zyloric 200 : 1 par jour](#) - Icaz LP 2.5 : 1 par jour - M  
 1 le matin - Aprovel 150 : 2 par jour - [Zyloric 200 : 1 par jour](#) - Ikaz LP 2.5 : 1 par jour - M  
 r le soir XATRAL LP : 1 cp/jour le soir [ZYLORIC 300 mg : 1 cp/jour le soir](#) CORDARONE 200 mg : 1  
 guinale Le traitement médical associé : [Zyprexa 5 mg/j](#), Lodoz 2.5 mg/j Histoire de la maladie :

## Etape 4



```

C:\Users\USER\AppData\Local\Unitex-GramLab\App>unitex.py
Remove old corpus-medical_snt if it exists so are you sure you want to execute UNITEX commands? (O/N): n
Execution aborted.

C:\Users\USER\AppData\Local\Unitex-GramLab\App>unitex.py
Remove old corpus-medical_snt if it exists so are you sure you want to execute UNITEX commands? (O/N): o
Le fichier spécifié est introuvable.
Normalizing corpus-medical.txt...

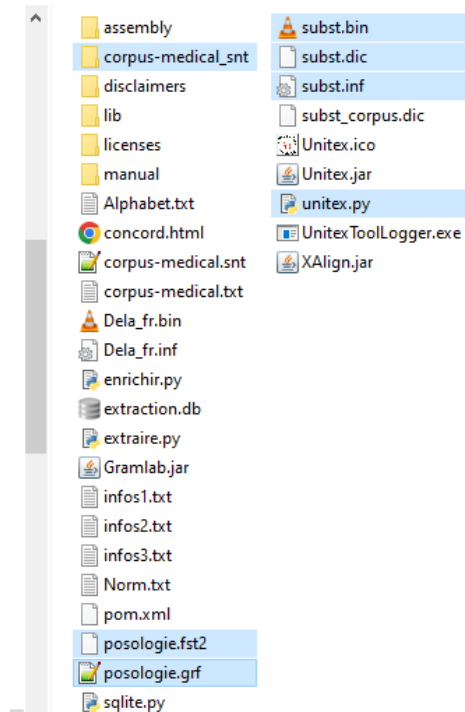
```

```

Looking for compound words...
First block...
Sorting and saving tag sequences...
Saving unknown words...
Done.
Compiling graph posologie.grf
Compilation has succeeded
Loading fst2...
Loading token list...
Loading morphological dictionaries...
Computing fst2 tags...
Loading dlf...
dlf: 10000 lines loaded...
Loading dlc...
Optimizing fst2 pattern tags...
Optimizing compound word dictionary...
Optimizing fst2...
Working...
100% done

1026 matches
12692 recognized units
(2.034% of the text is covered)
121695 exploration step
Done.
Loading concordance index...
Constructing concordance...
Done.

```





## Etape 5

```
C:\Users\USER\AppData\Local\Unitex-GramLab\App>sqlite.py
Usage: python sqlite.py [concord.html]

C:\Users\USER\AppData\Local\Unitex-GramLab\App>sqlite.py h
Error: Input file must be 'concord.html'

C:\Users\USER\AppData\Local\Unitex-GramLab\App>sqlite.py concord.html_
```

```
1001 | Fuflomedil 150 mg 1 matin et soir
1002 | Doliprane 500 mg 2 matin midi
1003 | Duphalac 2 c le matin
1004 | Furosemide 40 mg 2 le matin
1005 | Laroxyl 50 mg 1 le soir
1006 | Stilnox 10 mg 1 le soir
1007 | ATACAND 16 1/jour
1008 | TENORMINE 1001/jour
1009 | ZANIDIP 20 1 0 0
1010 | TEMESTA 2.5 2.5cp le soir
1011 | ALLOPURINOL 300 1/jour
1012 | doliprane 500 PIASCLEDINE 1/jour
1013 | TENORMINE 100, un le matin
1014 | HYPERIUM, un comprimé matin
1015 | glucose sur 24 heures
1016 | Seropram 20 1/j
1017 | Furosémide 20 2-0-0
1018 | augmentin: 1 g x 3 / jour
1019 | Furosémide 20 2-0-0
1020 | Seropram 20 1/j
1021 | solmucol :1 sachet x 3/ jour
1022 | FUROSEMIDE 20 mg : 2 cp le matin
1023 | INEXIUM 40 mg le soir
1024 | PROCTOLOG : 1 le matin
1025 | FORTIMEL : 2/j
1026 | TRANSIPEG : 2 sachets le matin
Inserted 1026 rows into the EXTRACTION table.
Data from concord.html successfully inserted into extraction.db.
```

Nom	Modifié le	Type	Taille
manuel	09/12/2023 00:06	Dossier de fichiers	
Alphabet.txt	06/08/2021 01:12	Document texte	1 Ko
concord.html	27/12/2023 20:48	Chrome HTML Do...	191 Ko
corpus-medical.snt	27/12/2023 20:48	Fichier SNT	3 257 Ko
corpus-medical.txt	08/12/2023 16:26	Document texte	1 749 Ko
Dela_fr.bin	06/08/2021 01:12	VLC media file (.bi...	4 435 Ko
Dela_fr.inf	06/08/2021 01:12	Informations de c...	305 Ko
enrichir.py	27/12/2023 20:43	Python File	4 Ko
extraction.db	27/12/2023 21:01	Fichier DB	52 Ko
extraire.py	26/12/2023 11:50	Python File	2 Ko

New Database

Open Database

Write Changes

Database Structure

Browse Data


































Edit Pragmas

Execute SQL

Table:

EXTRACTION

	ID	POSOLOGIE
	Filter	Filter
1003	1003	Dupnaia 2 c le matin
1004	1004	Furosemide 40 mg 2 le matin
1005	1005	Laroxyl 50 mg 1 le soir
1006	1006	Stilnox 10 mg 1 le soir
1007	1007	ATACAND 16 1/jour
1008	1008	TENORMINE 1001/jour
1009	1009	ZANIDIP 20 1 0 0
1010	1010	TEMESTA 2.5 2.5cp le soir
1011	1011	ALLOPURINOL 300 1/jour
1012	1012	doliprane 500 PIASCLEDINE 1/jour
1013	1013	TENORMINE 100, un le matin
1014	1014	HYPERIUM, un comprimé matin
1015	1015	glucose sur 24 heures
1016	1016	Seropram 20 1/j
1017	1017	Furosémide 20 2-0-0
1018	1018	augmentin: 1 g x 3 / jour
1019	1019	Furosémide 20 2-0-0
1020	1020	Seropram 20 1/j
1021	1021	solmucol :1 sachet x 3/ jour
1022	1022	FUROSEMIDE 20 mg : 2 cp le matin
1023	1023	INEXIUM 40 mg le soir
1024	1024	PROCTOLOG : 1 le matin
1025	1025	FORTIMEL : 2/j
1026	1026	TRANSIPEG : 2 sachets le matin

 assembly	 subst.bin
 corpus-medical_snt	 subst.dic
 disclaimers	 subst.inf
 lib	 subst_corpus.dic
 licenses	 Unitex.ico
 manual	 Unitex.jar
 Alphabet.txt	 unitex.py
 concord.html	 UnitexToolLogger.exe
 corpus-medical.snt	 XAlign.jar
 corpus-medical.txt	
 Dela_fr.bin	
 Dela_fr.inf	
 enrichir.py	
 extraction.db	
 extraire.py	
 Gramlab.jar	
 infos1.txt	
 infos2.txt	
 infos3.txt	
 Norm.txt	
 pom.xml	
 posologie.fst2	
 posologie.grf	
 sqlite.py	

---