Programmation appliquée à la chimie

Le cours “Programmation appliquée à la chimie” de bachelier en sciences chimiques (15 H cours et 15 H exercices, bloc2) utilise deux supports :

- Principalement, le présent wiki pour ses avantages techniques (coloration et indentation du code, recherche dans les pages, historique des modifications, …)
- Parfois, la plateforme moodle d'enseignement en ligne de l'UMONS, pour ses **autres** avantages techniques (authentification, devoirs, forum,…)

Si vous voulez comprendre la nécessité d'une formation informatique incluant l'initiation à l'algorithmique et la programmation, et l'apprentissage pratique via des projets, vous pouvez-consulter :

- Le Rapport de l’Académie des sciences - L’enseignement de l’informatique en France - Il est urgent de ne plus attendre (2013)
- Computer Science For All (President Obama in his 2016 State of the Union Address)

Notions de base (résumés, exemples de base,…)

- Présentation et principes de base de la programmation en Python, avec quelques applications + codes
- **Présentation et principes de base de la programmation en Python, avec quelques applications** (version wiki)
- Page avec de nombreuses informations sur Python (y compris pour l'apprentissage et l'installation). Modes recommandés :
  1. installation de la version 3 standard via python.org
  2. **installation de la distribution complète Anaconda (pour votre OS, en version 3 et 64 bits)**
     - User guide
     - ***Cf.* la page dédiée sur Anaconda**
  3. en salle informatique UMONS : utilisation de Python sous Ubuntu
  4. dans n'importe quel navigateur web, via le site https://repl.it/languages/python3
  5. utilisation en mode nomade de la plateforme Azure avec l'utilisation de vos codes UMONS
  6. via Colaboratory et l'utilisation d'un compte gmail, cette configuration permettant un partage et une édition en commun
- Notions fondamentales
- Les bases d'un interface graphique avec Tkinter
- Pièges à éviter

> Vous serez encouragé à l'utilisation et l'écriture de Jupyter notebooks, via l'utilisation de la distribution Anaconda, soit via la plateforme cloud Azure.
>
> *Cf* :
>
> - ce Tutoriel Jupyter en Jupyter (en construction)
> - cette application sur les données des éléments chimiques
> - A gallery of interesting Jupyter Notebooks

- ...

Algorithmes classiques

- Calcul de factorielles (pas à pas)
- Suite de Fibonacci (pas à pas)
- Manipulation de polynômes (pas à pas)
- Manipulations de matrices
- Algorithmes de tri
- Algorithmes de recherche
- Algorithmes sur entiers, où on parle d'atomes, de molécules, de nombres de nucléons !
- Algorithmes divers

Les librairies scientifiques essentielles

- Les bases de Matplotlib, une librairie pour réaliser des graphiques 2D
- Les bases de NumPy (tableaux numériques, algèbre linéaire, transformées de Fourier, nombres aléatoires)
- Les bases de SciPy : ajoute à NumPy des fonctionnalités mathématiques (intégration, optimisation, fonctions spéciales, interpolation,...)
- Pylab : permet de combiner simplement Matplotlib, NumPy et SciPy
- Pandas, pour l'analyse de données
- Scikit-learn, machine learning
- les bases de Plotly, une librairie pour réaliser des graphiques. plotly est interactif et s'intègre bien dans des navigateurs (*i.e.* Jupyter notebooks)
- les bases de Pygal, une autre librairie graphique, simple (cf. ce site)
- les bases de Altair, une autre librairie graphique, interactive

Notions intermédiaires et avancées

- Slices sur les listes
- Lire et écrire des fichiers de données csv (comma separated values)
- Notions avancées
- Trucs et astuces

Jupyter, IPython Notebooks et JupyterLab

Snippet de Wikipédia: ***Jupyter***

> **Jupyter** est une application web utilisée pour programmer dans plus de 40 langages de programmation, dont Python, Julia, Ruby, R, ou encore Scala. C'est un projet communautaire dont l'objectif est de développer des logiciels libres, des formats ouverts et des services pour l'informatique interactive. Jupyter est une évolution du projet IPython. Jupyter permet de

> réaliser des calepins ou notebooks, c'est-à-dire des programmes contenant à la fois du texte, simple ou enrichi typographiquement et sémantiquement grâce au langage à balises simplifié Markdown, et du code, lignes sources et résultats d'exécution. Ces calepins sont notamment utilisés en science des données pour explorer et analyser des données.

Creative Commons Attribution-Share Alike 4.0

- Jupyter : introduction, exemples, liens,...
- Les bases de Bokeh, une librairie Python de visualisations interactives pour des représentations dans des navigateurs web. Bokeh est particulièrement indiqué pour une utilisation dans les Jupyter notebooks, et s'installe aisément via Anaconda.

Applications

- Marche aléatoire 2D simple : simulation d'une chaîne polymère
- Flocon de Koch : simulation d'une fractale avec la librairie Turtle
- Traduction de l'ADN en séquence d'acides aminés (protéine) : utilisation d'un dictionnaire (type Python)
- Utilisation d'une "classe" pour des données de solvants chimiques
- Représentation du potentiel de Lennard-Jones, graphe élémentaire avec Matplotlib, et énergie de cohésion d'un cristal de gaz rare.
- Modélisation de la diffusion chimique dans un film : technique de différences finies, utilisation de Matplotlib
- Représentation de la distribution de vitesse de Maxwell-Boltzmann (avec et sans NumPy)
- Entropie de mélange pour un gaz ou liquide idéal
- Éléments et molécules : propriétés, masses molaires,...
- Représentation 3D du pH d'un acide en fonction d'un ajout de base et d'une dilution globale
- Analyse d'images
- L'attracteur de Lorenz : équations différentielles ordinaires et comportement chaotique
- OpenBabel et Jmol : format de description de molécules et visualisations
- Optimisation de la température caractéristique du diamant suivant le modèle d'Einstein (avec scipy, numpy, matplotlib)
- Test de Student : exemple technologique (avec scipy, numpy)
- Interrogation de la base de données géolocalisées OpenStreetMap
- Bioinformatique (manipulations de séquences ADN, ARN, protéines,...)
- Glossaire de chimie
- Mathématiques et nombres
- Épidémie du coronavirus COVID-19
- RDKit et molécules (représentation,...)

Données chimiques disponibles

- Données brutes non triées sur les éléments chimiques (pour usage après restructuration via du code python) :
  brute_data.zip
- Dataset de Kaggle sur les éléments chimiques
- Explore air pollution data European Environment Agency
- European Chemicals Agency

- …

Autres données

- Brent Oil Prices Daily historical Brent Oil Prices available on the U.S. Energy Information Admin
- Sources de données statistiques, données ouvertes (opendata), statistiques belges
  - OpenData : http://data.gov.be/fr
  - statbel : http://www.statbel.fgov.be/home_fr.asp
  - Belgostat online : http://www.nbb.be/app/cal/F/BelgoHome.htm
  - CREF : http://www.cref.be/Statistiques.htm
- statistiques européennes
  - eurostat : http://epp.eurostat.ec.europa.eu
- statistiques internationales
  - OECD : http://stats.oecd.org/wbos/
  - UNESCO : http://www.uis.unesco.org
  - United Nations : http://unstats.un.org/unsd/default.htm
  - World Trade Organization : http://www.wto.org/english/res_e/statis_e/its2005_e/its05_toc_e.htm
- Nations particulières
  - France :
    - https://www.data.gouv.fr/fr/
    - http://www.insee.fr/fr/home/home_page.asp
  - Luxembourg : http://www.statec.public.lu/en/
  - UK : http://www.statistics.gov.uk/
  - Hollande : http://www.cbs.nl/en/
  - Allemagne : http://www.destatis.de/e_home.htm
  - Italie : http://www.istat.it/English/index.htm
  - Espagne : http://www.ine.es/welcoing.htm
  - Suisse : http://www.bfs.admin.ch/bfs/portal/en/index.html
  - Canada : http://www.statcan.ca/start.html
  - UK : http://www.data.gov.uk
  - Australia : [http://data.gov.au/]]
  - USA :
    - http://www.fedstats.gov/
    - http://www.data.gov/ (et pages spécialisées, incluant la science et la recherche)
- Secteurs spécifiques
  - Economic Data freely available online
- Autres références :
  - http://www.statista.com/

Exemples de travaux d'étudiants

Ces travaux peuvent être entièrement originaux, ou se baser sur des éléments de code existants. + page à accès limité : lien intranet sur les travaux

- Jeu de la vie de Conway : automate cellulaire 2D (TkInter)
- Tableau périodique : tableau avec éléments cliquables pour obtenir les informations (y compris un fichier de données)

- Régression linéaire : entrée de couples, calcul et affichage de la droite de moindres carrés
- Ensemble de Mandelbrot : dessin d'une fractale
- Pavage de Penrose
- Courbe de Prédominance d'un Acide
- Création d'une grille et de configurations d'un système binaire modélisé
- Représentation de pH d'acides et de bases
- Représentation de molécules
- Calcul matriciel, avec l'interface Tk
- Conversion de températures, avec l'interface Tk
- pH et courbe de titrage, avec l'interface Tk
- Loi des gaz parfaits, avec l'interface Tk
- Tableau périodique : une autre version, avec l'interface Tk
- Traduction ADN-ARN-protéine, avec l'interface Tk
- Solubilité en fonction du pH et de la température, interface en ligne de commande et graphiques matplotlib
- Graphiques des pressions partielles de systèmes non-idéaux
- Vue 3D de l'électronégativité (tableau périodique)

Idées de travaux, projets

Consulter aussi cette [floss:python#applications_en_chimie|liste d'applications et librairies python en chimie]]

- Représentation de fonctions thermodynamiques de deux variables :
  - avec Matplotlib, en incluant des éléments supplémentaires
    - exemple : isothermes de van der Waals
  - Utiliser la librairie Mayavi (3D)
- Représentation des résolutions de Fourier pour la diffusion à 1D, 2D, 3D, en fonction du temps,...
- Résolution de problèmes numériques
  - intégration numérique
  - racines de polynômes, de fonctions générales
  - systèmes d'équations linéaires
  - optimisation de fonction (minimisations)
  - Approximations utilisant la formule de Stirling pour la factorielle (très utilisée en thermodynamique statistique), avec tabulation, représentations graphiques,...
- Interfacer Python et un tableur (par exemple pouvoir lire des données d'un fichier de tableur à partir d'un programme Python)
- Simulation en chimie :
  - degré de polymérisation comme l'article Software for Demonstration of Features of Chain Polymerization Processes
- Traitement d'images
- mesurer des temps de réaction (s'inspirer par exemple de fonctionnalités de ce programme)
- générer du son, de la vidéo
- Calcul de la constante de Madelung (interactions coulombiennes dans un cristal ionique)
- Représentation 3D du pH d'un acide en fonction d'un ajout de base et d'une dilution globale : cf. cet article : extension de l'exemple montré
- Échelle d'électronégativité, ou autre propriété atomique : représentation à 3D ou via des barres dont la taille est proportionnelle à la valeur, en suivant le schéma général du tableau périodique

- Illustrations graphiques des séries spectrales de l'hydrogène (barres, flèches, anneaux,...)
- Visualisations s'inspirant de sites comme ChemTube 3D, ChemEd DL,...
- Diagramme de Pourbaix à 3D : cf. cet article
- Outils de base en chimie comme sur le site http://fr.webqc.org/chemicaltools.php
- Multiéquilibre (mélange d'acides et bases conjuguées) : cf..cet article
- Utilisation de régressions non-linéaires pour traiter des données de réactions enzymatiques : cf. cet article
- Tableau périodique interactif via Ipython et Bokeh (cf. cette ref)
- Simulations de ségrégation, par inspiration de ces simulations, des travaux de Thomas Schelling, et sur base du programme Mesa (agent-based modelling)
- Équilibre liquide-vapeur à deux constituants, simulation du changement de phase (*cf.* cette simulation)
- Approximations de la fonction de Langevin utilisée pour décrire la magnétisation d'un matériau paramagnétique : représentation de la fonction et de ses approximations, des erreurs relatives,...
- graphiques polaires et cartésiens :
    - http://1ucasvb.tumblr.com/post/42881722643/the-familiar-trigonometric-functions-can-be
    - Illustration dynamique de séries de Fourier, comme ici
- Évolution de population suivant le modèle de la matrice de Leslie avec représentation de la pyramide des ages (*cf* ce lien)
- Créer des structures de données pour gérer les étiquettes de produits chimiques : pictogrammes, mentions de danger (Hxxx et EUHxxx), de prudence (Pxxx), codes produits,...
- Rotamères : étude simplifiée et représentation graphique (*cf.* cet exemple)
- Décomposition spinodale (modèle de Cahn-Hilliard), (*cf.* cet exemple)
- Programme basé sur une de ces librairies :
    - pyEQL, librairie pour gérer des solutions aqueuses d'électrolytes
    - chemlab, librairies incluant la visualisation et la manipulation de données sur les structures chimiques
    - chempy, librairie pour résoudre des problèmes de chimie physique, chimie analytique,...
- ChEMBL base de donnée chimique
- ChemSpiPy pour accéder à la base de donnée ChemSpider
- Wikipedia chemical structure explorer, avec données exploitables via GitHub
- Charte des nucleides, et diverses représentations des instabilités. cf. Live Chart of Nuclides et les données sources indiquées
- **collisions de particules (sphères ou disques)**
    - Molecular Dynamics Simulation of Hard Spheres → description générale du problème avec des solutions basées sur les particules ou les événements (hard disc simulation, dynamique moléculaire) :
    - documents de Werner Krauth et al.
        - Main Page - Werner KRAUTH
        - Lecture 2: Hard disks: from Classical Mechanics to Statistical Mechanics - Hard disks: From Classical Mechanics to Statistical Mechanics | Coursera
        - Bernard Krauth 2012 - Werner KRAUTH
    - Event-driven
        - Event-Driven Simulation
        - numberset/HardSphereSim: Event driven simulation of hard spheres in python
        - rajeshrinet/pyedmd: Event-Driven Molecular dynamics simulations
        - robeme/HardSphereSim: Event driven simulation of hard spheres in python
    - Autres simulations
        - ideal-gas-simulation/simulation.py at master · labay11/ideal-gas-simulation
        - gas.py

- Exploitation de données chimiques de wikidata.org query (> 20000 composés),

```
PREFIX wdt: <http://www.wikidata.org/prop/direct/>
  SELECT ?compound WHERE {
    ?compound wdt:P31 wd:Q11173 .
  }
```

- chaines de Markov sur des séquences peptidiques
- Simulation électrochimique comme dans cet article
- Splitting en RMN du proton, comme cet article
- Catalyse biochimique (Michaelis-Menten) : représentation et calcul des constantes comme dans cet article
- Simulation de l'expansion irréversible d'un gaz mono-atomique, comme dans cet article
- représentation de liaisons, densités de charge, comme dans cet article
- simulation du modèle d'Ising : http://rajeshrinet.github.io/blog/2014/ising-model/
- visualisation comme dans l'article Visualization of Buffer Capacity with 3-D “Topo” Surfaces: Buffer Ridges, Equivalence Point Canyons and Dilution Ramps ou 3-D Topo Surface Visualization of Acid–Base Species Distributions: Corner Buttes, Corner Pits, Curving Ridge Crests, and Dilution Plains
- sur base de A Python Program for Solving Schrödinger’s Equation in Undergraduate Physical Chemistry
- programmes en version python analogues aux applications en chimie de l'article The Development of Computational Thinking in a High School Chemistry Course Paul S. Matsumoto and Jiankang Cao, J. Chem. Educ., 2017, 94 (9), pp 1217–1224 DOI: 10.1021/acs.jchemed.6b00973
- programmes inspirés de Teaching the Growth, Ripening, and Agglomeration of Nanostructures in Computer Experiments Jan Philipp Meyburg and Detlef Diesing, J. Chem. Educ., 2017, 94 (9), pp 1225–1231 DOI: 10.1021/acs.jchemed.6b01008
- Créer un nuage de tag | mots ou word cloud à partir de textes publications websites de chimie (utiliser scikit-learn et/ou du “web scrapping”,...). Liens : Word Cloud, PyTagCloud, ref1, ref2, ref3,...
- représentations d'orbitales électroniques 1D ou 3D.
  - *Cf.* par exemple cet article
  - http://anguslowe.me/physics/orbitals/
- Colorimétrie : Color Space Mathematical Modeling Using Microsoft Excel M. J. Delgado-González, Y. Carmona-Jiménez, M. C. Rodríguez-Dodero, and M. V. García-Moreno, J. Chem. Educ., 2018, 95 (10), pp 1885–1889 DOI: 10.1021/acs.jchemed.7b00681
- Énergie électrique et émission de CO2, en continu : https://www.electricitymap.org
- Données diverses (France) : http://data.cquest.org/
- Visualisation de la taille d'effet : cf. https://rpsychologist.com/d3/cohend/
- simulation du Problème de Monty Hall (+ Monty Hall problem, Paradoxe des prisonniers)
- ...

Voir aussi la sélection de codes sources ActiveState avec de nombreuses applications scientifiques.

Voir aussi ces librairies.

Références générales

- Coder proprement Robert C. Martin 2019, Eyrolles

- 21 Python Mini Projects With Code - Get a Speed Boost In Your Python Journey By Building These Amazing Projects Abhay Parashar, Medium, 04/01/2021

From:
https://dvillers.umons.ac.be/wiki/ - **Didier Villers, UMONS - wiki**

Permanent link:
https://dvillers.umons.ac.be/wiki/teaching:progappchim:start?rev=1612194660

Last update: **2021/02/01 16:51**