

M2 AMIS – Module Projet

Franck.Quessette@uvsq.fr & Sandrine.vial@uvsq.fr

11 octobre 2022

- Le module projet se déroule par groupes de cinq ou six étudiants. Les groupes sont donnés à la fin de ce document.
- Chaque groupe doit s'auto-organiser. Cette auto-organisation fait partie du projet.
- Ce module est **obligatoire** et il n'y a pas de seconde session.

Table des matières

1	Sujet du projet	1
1.1	Étape 1 : récupération des données	1
1.2	Étape 2 : parser les données	2
1.3	Étape 3 : définir une structure de données	2
1.4	Étape 4 : utilisation de l'algo de McKay	2
1.5	Étape 5 : intégration	2
2	Sites de référence	2
3	Calendrier	3
4	Liste des étudiants par groupe	4

1 Sujet du projet

Le but est d'écrire un programme qui permet de tester si des molécules sont de structures identiques.

1.1 Étape 1 : récupération des données

Il faut écrire un programme qui récupère le ou les fichiers contenant les molécules et stocke ce ou ces fichiers en local sur l'ordinateur. L'adresse web de récupération doit être prise en argument sur la ligne de commande.

1.2 Étape 2 : parser les données

Il faut parser les fichiers récupérés et créer un fichier par molécule, le fichier d'une molécule doit être nommé. Vous devez définir le format de fichier pour une molécule et les infos qu'il doit contenir. Vous pouvez soit utiliser un format déjà existant soit définir le vôtre. Vous devez définir comment nommer chaque fichier.

1.3 Étape 3 : définir une structure de données

Vous devez définir une structure de données permettant de stocker une molécule. Vous avez le choix du langage de programmation. Le choix de la structure de données est très lié à l'utilisation que vous allez en faire à l'étape 4.

1.4 Étape 4 : utilisation de l'algo de McKay

Le problème de l'isomorphisme de graphes n'est pas décidé comme étant dans P ou dans NP. Il existe néanmoins un algorithme très efficace dû à Brendan McKay.

Vous devez expliquer en quoi le problème d'isomorphisme de graphes permet de tester l'équivalence structurelle de molécules.

Vous devez utiliser la bibliothèque nauty pour décider si deux molécules sont identiques ou pas

1.5 Étape 5 : intégration

Vous devez intégrer les différentes étapes (Makefile ou autre). La présentation des paires de molécules identiques doit permettre d'être facilement vérifié en utilisant le site de chebi.

2 Sites de référence

La liste des molécules à étudier est celle qui est sur le site de chebi :

— <https://www.ebi.ac.uk/chebi/>.

Vous devez lire la documentation et comprendre comment récupérer les ou les fichiers contenant les molécules :

— <https://docs.google.com/document/d/11G6SmTtQRQYFT7l9h5K0faUHiAaekcLe0weMOOTIpME/edit>.

Exemple d'affichage d'une molécule directement avec son numéro chebi :

— <https://www.ebi.ac.uk/chebi/searchId.do?chebiId=CHEBI:27732> ;
— <https://www.ebi.ac.uk/chebi/searchId.do?chebiId=CHEBI:6809> ;
— <https://www.ebi.ac.uk/chebi/searchId.do?chebiId=CHEBI:27808>.

La page wikipedia sur l'isomorphisme de graphes :

— https://fr.wikipedia.org/wiki/Isomorphisme_de_graphes

Le site de nauty et Traces de McKay et Piperno :

— <http://users.cecs.anu.edu.au/~bdm/nauty/>

Le manuel d'utilisation de nauty et Traces :

— <http://users.cecs.anu.edu.au/~bdm/nauty/nug27.pdf>

3 Calendrier

Les différents rendez-vous du projet avec les enseignants :

lundi 03 octobre 2022	Présentation du projet.
vendredi 21 octobre 2021	Séance de suivi pour éclaircir le sujet.
vendredi 18 novembre 2022	Présentation orale de 20 minutes sur l'organisation du travail et montrant la compréhension des points techniques :
13h50 – 14h10	Groupe A
14h10 – 14h30	Groupe B
14h30 – 15h50	Groupe C
	Le diaporama de la présentation doit être envoyé aux enseignants au moment de la soutenance.
jeudi 08 décembre 2023	Séance de suivi.
xx décembre 2021	Présentation orale de 20 minutes sur les formats de fichier, le/les langages de programmation et les structures de données proposés :
13h50 – 14h10	Groupe B
14h10 – 14h30	Groupe C
14h30 – 15h50	Groupe A
	Le diaporama de la présentation doit être envoyé aux enseignants au moment de la soutenance.
début janvier 2022	Séance de suivi.
mi-fin janvier 2022	Remise du rapport et soutenance finale de 30 minutes :
13h50 – 14h20	Groupe C
14h20 – 14h50	Groupe A
14h50 – 15h20	Groupe B
	Le rapport doit être envoyé aux enseignants la veille de la soutenance. Le diaporama de la présentation doit être envoyé aux enseignants au moment de la soutenance.

4 Liste des étudiants par groupe

	Gr.	NOM	Prénom	mail	Num. étu.
1	A	ABDALLAH	Farouk	farouk.abdallah@ens.uvsq.fr	22209367
2	A	AKROUNE	Feriel	feriel.akroune@ens.uvsq.fr	21807878
3	A	BOUTOUILI	Djillali	djillali.boutouili@ens.uvsq.fr	22107265
4	A	CAMBRESY	Florian	florian.cambresy@ens.uvsq.fr	21803824
5	A	HELAIN	Chloe	chloe.helain@ens.uvsq.fr	21800801
6	A	OUHOCINE	Sarah	sarah.ouhocine@ens.uvsq.fr	22107541
1	B	ABERKANE	Walid	walid.aberkane@ens.uvsq.fr	22106783
2	B	BARBIER	Yann	yann.barbier@ens.uvsq.fr	21800844
3	B	CHALAUD	Jean-Christophe	jean-christophe.chalaud@ens.uvsq.fr	21801285
4	B	GOUY	Emeric	emeric.gouy@ens.uvsq.fr	21800030
5	B	RAZAFIMBELO	Fanoa	fanoa.razafigimbelo@ens.uvsq.fr	22105665
1	C	BENDAHGANE	Nassim	nassim.bendahgane@ens.uvsq.fr	22208003
2	C	BOUSKRA	Nazim	nazim.bouskra@ens.uvsq.fr	22106254
3	C	MANSOURA	Mohamed Amine	mohamed-amine.mansoura@ens.uvsq.fr	21917483
4	C	SOURSOU	Adrien	adrien.soursou@ens.uvsq.fr	21601601
5	C	ZENS	Coralie	coralie.zens@ens.uvsq.fr	21800138
	Abs	ABDALLAH	Farouk	farouk.abdallah@ens.uvsq.fr	22209367
	Abs	TOUZARI	Maya	maya.touzari@ens.uvsq.fr	22206351