

# M2 AMIS – Module Projet

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

14 novembre 2025

- 
- 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

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

#### 1.1 Algo de base

##### 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.

##### 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.

##### 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.

## 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 du 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

## 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.

### 1.2 Algo avancé

Vous devez proposer une ou plusieurs mesures de distance entre deux molécules.

Pour chacune des distances que vous jugerez pertinente vous ferez un/des clusterings.

L'analyse des résultats obtenus est une part importante de ce projet.

## 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 le ou les fichiers contenant les molécules :

- <https://docs.google.com/document/d/11G6SmTtQRQYFT719h5K0faUHiAaekcLeOweMOOTIpME/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](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 :

vendredi 14 novembre 2025	Présentation du projet.
vendredi 5/12 décembre 2025	Séance de suivi pour éclaircir le sujet.
mi janvier 2026  13h50 – 14h10 14h20 – 14h40 14h50 – 15h10 15h20 – 15h40	Présentation orale de 20 minutes sur l'organisation du travail et montrant la compréhension des points techniques : <b>Groupe A</b> <b>Groupe B</b> <b>Groupe C</b> <b>Groupe D</b> Le diaporama de la présentation doit être envoyé aux enseignants au moment de la soutenance.
début février 2026  13h50 – 14h20 14h30 – 15h00 15h10 – 15h40 15h50 – 16h20	Remise du rapport et soutenance finale de 30 minutes : <b>Groupe D</b> <b>Groupe C</b> <b>Groupe B</b> <b>Groupe A</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	AARRIS	YANIS	yanis.arris@ens.uvsq.fr	22402904
2	BOUAZIZ	AMIRA	amira.bouaziz@ens.uvsq.fr	22504378
3	CHAUMONT	NICOLAS	nicolas.chaumont@ens.uvsq.fr	21914689
4	CHEDOZEAU	MATHIEU	mathieu.chedozeau@ens.uvsq.fr	22102055
5	HAMOUDI	NABIL	nabil.hamoudi@ens.uvsq.fr	22006316
6	SRAIRI	AISSA WAIL	aissa-wail.srairi@ens.uvsq.fr	22406233
1	DUCROS	ALEXANDRE	alexandre.ducros@ens.uvsq.fr	22404016
2	HIMEUR	WAIL	wail.himeur@ens.uvsq.fr	22506340
3	KORELSKAIA	KIRA	kira.korelskaia@ens.uvsq.fr	22500346
4	MACHE	ETHAN	ethan.mache@ens.uvsq.fr	22101464
5	THIBAUDEAU	CYRIAC	cyriac.thibaudeau@ens.uvsq.fr	22104322
6	TREBERN	EWAN	ewan.trebern@ens.uvsq.fr	21916443
1	BEJJANI	ALAIN	alain.bejjani@ens.uvsq.fr	22405154
2	DUPUY	RAPHAEL	raphael.dupuy@ens.uvsq.fr	22101595
3	NGOUFACK ZAGOUUM	CARNOT SORENSEN	carnot-sorensen.ngoufack-zagouum@ens.uvsq.fr	22309034
4	RICHARD	CLEMENT	clement.richard@ens.uvsq.fr	21917563
5	TEHAMMI	SAFIA	safia.tehami@ens.uvsq.fr	22500828