Master Informatique - M1 - UE Complexité Chapitre 1 : Complexité, Présentation et Introduction

Philippe Jégou

Laboratoire d'Informatique et Systèmes - LIS - UMR CNRS 7020 Équipe COALA - COntraintes, ALgorithmes et Applications Campus de Saint-Jérôme

Département Informatique et Interactions
Faculté des Sciences
Université d'Aix-Marseille

philippe.jegou@univ-amu.fr

4 septembre 2020



Plan

- Organisation de l'UE
- Évaluation
- 3 Sur la matière enseignée
- Plan du cours

Plan

- Organisation de l'UE
- <u>Évaluation</u>
- 3 Sur la matière enseignée
- Plan du cours

Organisation de l'UE

Séances :

- 20h de Cours en 10 séances de 2h :
 - 5 premières séances : Philippe Jégou
 - 5 dernières séances : Antonio E. Porreca
- 20h de Travaux Dirigés en 10 séances de 2h avec 3 groupes et donc 3 intervenants :
 - Antonio F. Porreca
 - Nicolas Prcovic
 - Raphaël Sturgis
- 14h de Travaux Pratiques en 7 séances de 2h avec 4 groupes et donc 4 intervenants :
 - François Hamonic
 - Antonio E. Porreca
 - Nicolas Prcovic
 - Raphaël Sturgis



Communication : Où, quoi, commment ?

Site internet du module

 C'est ici que ça se passe : http://ametice.univ-amu.fr/

Communication : Où, quoi, commment ?

Site internet du module

 C'est ici que ça se passe : http://ametice.univ-amu.fr/

Vous y trouverez:

- les informations sur l'organisation du module : dates, changements, annonces, etc.
- les documents mis à votre disposition : diapositives des cours, sujets de TD, de TP, etc.
- un forum pour échanger, communiquer entre vous, avec les enseignants, etc.

• D'ailleurs pour communiquer : utilisez le forum !!!

- et si un message à un enseignant doit absolument passer hors forum :
 - utilisez pour l'envoi votre adresse AMU
 - l'objet de votre message devra commencer par "[M1-Info-Complexité]"

Plan

- Organisation de l'UE
- Évaluation
- Sur la matière enseignée
- Plan du cours

UE Complexité : 6 crédits soit 20 % du 1er semestre !

UE Complexité : 6 crédits soit 20 % du 1er semestre !

MCC ?

Les fameuses Modalités de Contrôle des Connaissances

- Session 1: $0.75 \times max((CC + (2 \times ET))/3, ET) + 0.25 \times TP$
 - ET : examen terminal en amphi
 - CC : contrôle continu a priori avec un partiel à la "mi-temps"
 - TP : note de TP (sous la forme de petits projets réalisés en groupes)

UE Complexité : 6 crédits soit 20 % du 1er semestre !

MCC ?

Les fameuses Modalités de Contrôle des Connaissances

- Session 1: $0.75 \times max((CC + (2 \times ET))/3, ET) + 0.25 \times TP$
 - ET: examen terminal en amphi
 - CC : contrôle continu a priori avec un partiel à la "mi-temps"
 - TP : note de TP (sous la forme de petits projets réalisés en groupes)
- Session 2 : $max(ET, (0, 75 \times ET) + 0, 25 \times TP)$



UE Complexité : 6 crédits soit 20 % du 1er semestre !

MCC ?

Les fameuses Modalités de Contrôle des Connaissances

- Session 1: $0.75 \times max((CC + (2 \times ET))/3, ET) + 0.25 \times TP$
 - ET: examen terminal en amphi
 - CC : contrôle continu a priori avec un partiel à la "mi-temps"
 - TP : note de TP (sous la forme de petits projets réalisés en groupes)
- Session 2 : $max(ET, (0, 75 \times ET) + 0, 25 \times TP)$
- Attention !

Sous réserve de modifications en relation avec la crise sanitaire



Plan

- Organisation de l'UE
- Évaluation
- 3 Sur la matière enseignée
- Plan du cours

Connaissances requises pour démarrer l'UE :

- Algorithmique :
 - conception d'algorithmes
 - analyse de la complexité (surtout pire des cas et "grand O")
 - algorithmique des graphes

Connaissances requises pour démarrer l'UE :

- Algorithmique :
 - conception d'algorithmes
 - analyse de la complexité (surtout pire des cas et "grand O")
 - algorithmique des graphes
- Théorie de Langages Formels :
 - notions de base : alphabet, mots, langages, etc.
 - automates finis
 - machines de Turing

Connaissances requises pour démarrer l'UE :

- Algorithmique :
 - conception d'algorithmes
 - analyse de la complexité (surtout pire des cas et "grand O")
 - algorithmique des graphes
- Théorie de Langages Formels :
 - notions de base : alphabet, mots, langages, etc.
 - automates finis
 - machines de Turing
- Logique des Propositions
 - notions de base : variables, clauses, interprétation, CNF
 - traitement de formules : Quine, DPLL, Résolution

Connaissances requises pour démarrer l'UE :

- Algorithmique :
 - conception d'algorithmes
 - analyse de la complexité (surtout pire des cas et "grand O")
 - algorithmique des graphes
- Théorie de Langages Formels :
 - notions de base : alphabet, mots, langages, etc.
 - automates finis
 - machines de Turing
- Logique des Propositions
 - notions de base : variables, clauses, interprétation, CNF
 - traitement de formules : Quine, DPLL, Résolution

En résumé : tout ce que vous avez adoré en licence !...

Et en plus, des rappels seront réalisés



Quelques ouvrages de référence :

T. Cormen, C. Leiserson, R. Rivest, et C. Stein.
 Algorithmique. Dunod, 2010.

Ouvrage de référence sur l'Algorithmique : vous connaissez déjà

Quelques ouvrages de référence :

- T. Cormen, C. Leiserson, R. Rivest, et C. Stein. Algorithmique. Dunod, 2010.
 - Ouvrage de référence sur l'Algorithmique : vous connaissez déjà
- M. R. Garey and D. S. Johnson. Computers and Intractability A Guide to the Theory of NP-Completeness. Freeman and Co., 1979.
 - Livre de référence sur la Théorie de la Complexité : Appelé "Le Garey et Johnson" !
 - Fin août 2020 : 67643 citations selon Google Scholar !
 Soit plus de 30 fois la référence la plus citée de...

Quelques ouvrages de référence :

- T. Cormen, C. Leiserson, R. Rivest, et C. Stein. Algorithmique. Dunod, 2010.
 - Ouvrage de référence sur l'Algorithmique : vous connaissez déjà
- M. R. Garey and D. S. Johnson. Computers and Intractability A Guide to the Theory of NP-Completeness. Freeman and Co., 1979.
 - Livre de référence sur la Théorie de la Complexité : Appelé "Le Garey et Johnson" !
 - Fin août 2020 : 67643 citations selon Google Scholar !
 Soit plus de 30 fois la référence la plus citée de...
- P. Wolper. Introduction à la Calculabilité. Dunod, 2006.
 Facile d'accès et une partie reprend (en français!) ce que nous allons raconter du "Garey et Johnson"

Quelques ouvrages de référence :

- T. Cormen, C. Leiserson, R. Rivest, et C. Stein. Algorithmique. Dunod, 2010.
 - Ouvrage de référence sur l'Algorithmique : vous connaissez déjà
- M. R. Garey and D. S. Johnson. Computers and Intractability A Guide to the Theory of NP-Completeness. Freeman and Co., 1979.
 - Livre de référence sur la Théorie de la Complexité : Appelé "Le Garey et Johnson" !
 - Fin août 2020 : 67643 citations selon Google Scholar !
 Soit plus de 30 fois la référence la plus citée de...
- P. Wolper. Introduction à la Calculabilité. Dunod, 2006.
 Facile d'accès et une partie reprend (en français!) ce que nous allons raconter du "Garey et Johnson"
- et plein d'autres pour aller plus loin ; la liste sera sur Ametice



- Va bien au-delà de la complexité des algorithmes
 - Analyse de la complexité des algorithmes :
 Estimer le temps (théorique) d'exécution d'un algorithme
 (pour estimer le temps effectif d'exécution du programme associé)

- Va bien au-delà de la complexité des algorithmes
 - Analyse de la complexité des algorithmes :
 Estimer le temps (théorique) d'exécution d'un algorithme
 (pour estimer le temps effectif d'exécution du programme associé)
 - lci : saut conceptuel car c'est la difficulté des problèmes qui est analysée :
 Quel est le temps (théorique) nécessaire pour résoudre un problème ?
 En d'autres termes : Quelle est la difficulté d'un problème ?

- Va bien au-delà de la complexité des algorithmes
 - Analyse de la complexité des algorithmes :
 Estimer le temps (théorique) d'exécution d'un algorithme
 (pour estimer le temps effectif d'exécution du programme associé)
 - lci : saut conceptuel car c'est la difficulté des problèmes qui est analysée :
 Quel est le temps (théorique) nécessaire pour résoudre un problème ?
 En d'autres termes : Quelle est la difficulté d'un problème ?
- Des questions fondamentales :
 - Beaucoup de travaux (théoriques) réalisés depuis (plus de) 50 ans
 - Et toujours une interrogation fondamentale : est-ce que $P \neq NP$? La fameuse question à 1 million de \$



- Va bien au-delà de la complexité des algorithmes
 - Analyse de la complexité des algorithmes :
 Estimer le temps (théorique) d'exécution d'un algorithme
 (pour estimer le temps effectif d'exécution du programme associé)
 - lci : saut conceptuel car c'est la difficulté des problèmes qui est analysée :
 Quel est le temps (théorique) nécessaire pour résoudre un problème ?
 En d'autres termes : Quelle est la difficulté d'un problème ?
- Des questions fondamentales :
 - Beaucoup de travaux (théoriques) réalisés depuis (plus de) 50 ans
 - Et toujours une interrogation fondamentale : est-ce que $P \neq NP$? La fameuse question à 1 million de \$
- Et c'est pas toujours simple sur le plan pratique !

Un cas d'étude très concret qui pourrait vous concerner (dans 2 ans)

Un cas d'étude très concret qui pourrait vous concerner (dans 2 ans)

Une startup doit realiser le projet qui a motivé sa création :
 Un système de calcul de communautés dans un réseau social

Un cas d'étude très concret qui pourrait vous concerner (dans 2 ans)

- Une startup doit realiser le projet qui a motivé sa création :
 Un système de calcul de communautés dans un réseau social
- Communautés dans ce réseau social ?
 Plusieurs types de communautés ici :
 - Communautés très élargies
 - Communautés élargies
 - Communautés fortes

Types de communautés : définis par les liens existants entre membres du réseau

Liens entre membres : existence d'envois de messages



Communautés très élargies :

Communautés très élargies :

- ullet X et Y sont dans la même communauté très élargie :
 - si X a envoyé un message à Y (représenté par $X \to Y$)
 - ou si Y a envoyé un message à X
 - donc si X et Y se sont échangé au moins un message

Communautés très élargies :

- X et Y sont dans la même communauté très élargie :
 - ullet si X a envoyé un message à Y (représenté par X o Y)
 - ou si Y a envoyé un message à X
 - donc si X et Y se sont échangé au moins un message

mais aussi:

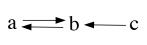
X et Y sont dans la même communauté :
 si on a un Z qui est dans la même communauté que X et que Y
 (X et Y peuvent n'avoir jamais échangé de message)

Communautés très élargies :

- X et Y sont dans la même communauté très élargie :
 - si X a envoyé un message à Y (représenté par $X \to Y$)
 - \bullet ou si Y a envoyé un message à X
 - donc si X et Y se sont échangé au moins un message

mais aussi :

X et Y sont dans la même communauté :
 si on a un Z qui est dans la même communauté que X et que Y
 (X et Y peuvent n'avoir jamais échangé de message)







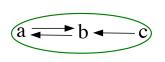
Quelles sont les communautés très élargies ?

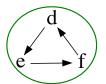
Communautés très élargies :

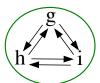
- X et Y sont dans la même communauté très élargie :
 - si X a envoyé un message à Y (représenté par $X \to Y$)
 - \bullet ou si Y a envoyé un message à X
 - donc si X et Y se sont échangé au moins un message

mais aussi :

X et Y sont dans la même communauté :
 si on a un Z qui est dans la même communauté que X et que Y (là, X et Y peuvent n'avoir jamais échangé de message)







3 communautés très élargies : $\{a, b, c\}$, $\{d, e, f\}$ et $\{g, h, i\}$

Communautés élargies :

Communautés élargies :

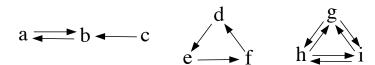
- X et Y sont dans la même communauté élargie :
 - ullet s'il existe une séquence de messages partant de X qui arrive jusqu'à Y, et
 - s'il existe une séquence de messages partant de Y qui arrive jusqu'à X

Communautés élargies :

- X et Y sont dans la même communauté élargie :
 - ullet s'il existe une séquence de messages partant de X qui arrive jusqu'à Y, et
 - s'il existe une séquence de messages partant de Y qui arrive jusqu'à X et ainsi :
- tous les membres Z du réseau figurant sur cette séquence d'envois sont dans la même communauté que X et Y

Communautés élargies :

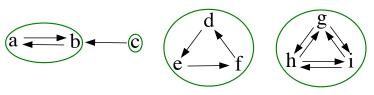
- X et Y sont dans la même communauté élargie :
 - ullet s'il existe une séquence de messages partant de X qui arrive jusqu'à Y, et
 - ullet s'il existe une séquence de messages partant de Y qui arrive jusqu'à X et ainsi :
- tous les membres Z du réseau figurant sur cette séquence d'envois sont dans la même communauté que X et Y



Quelles sont les communauté élargies ?

Communautés élargies :

- X et Y sont dans la même communauté élargie :
 - ullet s'il existe une séquence messages partant de X et qui arrive jusqu'à Y, et
 - s'il existe une séquence messages partant de Y et qui arrive jusqu'à X
 et ainsi :
- tous les membres Z du réseau figurant sur cette séquence d'envois sont dans la même communauté que X et Y



4 communautés élargies : $\{a,b\}$, $\{c\}$, $\{d,e,f\}$ et $\{g,h,i\}$

Communautés fortes :

Communautés fortes :

- X et Y sont dans la même communauté forte :
 - ullet si X a envoyé un message à Y, et
 - ullet si Y a envoyé un message à X

Communautés fortes :

- X et Y sont dans la même communauté forte :
 - si X a envoyé un message à Y, et
 - si Y a envoyé un message à X

et ainsi:

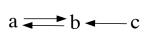
• tous les membres du réseau d'une même communauté forte se sont échangés deux à deux des messages

Communautés fortes :

- X et Y sont dans la même communauté forte :
 - ullet si X a envoyé un message à Y, et
 - \bullet si Y a envoyé un message à X

et ainsi :

• tous les membres du réseau d'une même communauté forte se sont échangés deux à deux des messages







Quelles sont les communautés fortes ?

Communautés fortes :

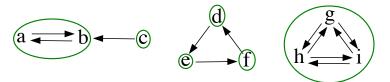
- X et Y sont dans la même communauté forte :
 - si X a envoyé un message à Y,

et

si Y a envoyé un message à X

et ainsi :

 tous les membres du réseau d'une même communauté forte se sont échangés deux à deux des messages



6 communautés fortes : $\{a,b\}$, $\{c\}$, $\{d\}$, $\{e\}$, $\{f\}$ et $\{g,h,i\}$

- Retour sur le projet de la startup : pour chaque type de communauté :
 - Communautés très élargies
 - Communautés élargies
 - Communautés fortes

il faut identifier la plus grande communauté

(les motivations vont rester secrètes car secret industriel...)

- Retour sur le projet de la startup : pour chaque type de communauté :
 - Communautés très élargies
 - Communautés élargies
 - Communautés fortes
 - il faut identifier la plus grande communauté (les motivations vont rester secrètes car secret industriel...)
- Un investisseur mettra-t-il des moyens ?

- Retour sur le projet de la startup : pour chaque type de communauté :
 - Communautés très élargies
 - Communautés élargies
 - Communautés fortes

il faut identifier la plus grande communauté (les motivations vont rester secrètes car secret industriel...)

- Un investisseur mettra-t-il des moyens ?
 - Communautés très élargies : NON !
 - → calcul des **composantes connexes** d'un graphe
 - \rightarrow facile et en $\Theta(n+m)$ (vu en Licence)

- Retour sur le projet de la startup : pour chaque type de communauté :
 - Communautés très élargies
 - Communautés élargies
 - Communautés fortes

il faut identifier la plus grande communauté (les motivations vont rester secrètes car secret industriel...)

- Un investisseur mettra-t-il des moyens ?
 - Communautés très élargies : NON !
 - → calcul des **composantes connexes** d'un graphe
 - \rightarrow facile et en $\Theta(n+m)$ (vu en Licence)
 - Communautés élargies : NON !
 - → calcul des composantes fortement connexes d'un graphe orienté
 - \rightarrow subtil et en $\Theta(n+m)$ (vu en Licence)

- Retour sur le projet de la startup : pour chaque type de communauté :
 - Communautés très élargies
 - Communautés élargies
 - Communautés fortes

il faut identifier la plus grande communauté (les motivations vont rester secrètes car secret industriel...)

- Un investisseur mettra-t-il des moyens ?
 - Communautés très élargies : NON !
 - → calcul des **composantes connexes** d'un graphe
 - \rightarrow facile et en $\Theta(n+m)$ (vu en Licence)
 - Communautés élargies : NON !
 - → calcul des composantes fortement connexes d'un graphe orienté
 - \rightarrow subtil et en $\Theta(n+m)$ (vu en Licence)
 - Communautés fortes ????



Calculer la plus grande communauté forte ? C'est :

- pas vu en Licence
- résoudre le problème de la clique maximum (graphe non orienté)
- un problème conjecturé de complexité exponentielle...

Calculer la plus grande communauté forte ? C'est :

- pas vu en Licence
- résoudre le problème de la clique maximum (graphe non orienté)
- un problème conjecturé de complexité exponentielle...

Si la startup découvre un algorithme polynomial :

elle gagnera tout de suite 1 million de dollars (prix CMI)

Calculer la plus grande communauté forte ? C'est :

- pas vu en Licence
- résoudre le problème de la clique maximum (graphe non orienté)
- un problème conjecturé de complexité exponentielle...

- elle gagnera tout de suite 1 million de dollars (prix CMI)
- et ses membres :

Calculer la plus grande communauté forte ? C'est :

- pas vu en Licence
- résoudre le problème de la clique maximum (graphe non orienté)
- un problème conjecturé de complexité exponentielle...

- elle gagnera tout de suite 1 million de dollars (prix CMI)
- et ses membres :
 - un Turing Award, une médaille Fields (s'ils ont moins de 40 ans), un prix Nobel d'Économie, le prix Gödel...

Calculer la plus grande communauté forte ? C'est :

- pas vu en Licence
- résoudre le problème de la clique maximum (graphe non orienté)
- un problème conjecturé de complexité exponentielle...

- elle gagnera tout de suite 1 million de dollars (prix CMI)
- et ses membres :
 - un Turing Award, une médaille Fields (s'ils ont moins de 40 ans), un prix Nobel d'Économie, le prix Gödel...
 - un poste de professeur au MIT, ou à Standord, Princeton, Harvard, Oxford, Cambridge, etc... ou éventuellement à AMU

Calculer la plus grande communauté forte ? C'est :

- pas vu en Licence
- résoudre le problème de la clique maximum (graphe non orienté)
- un problème conjecturé de complexité exponentielle...

- elle gagnera tout de suite 1 million de dollars (prix CMI)
- et ses membres :
 - un Turing Award, une médaille Fields (s'ils ont moins de 40 ans), un prix Nobel d'Économie, le prix Gödel...
 - un poste de professeur au MIT, ou à Standord, Princeton, Harvard, Oxford, Cambridge, etc... ou éventuellement à AMU
 - et mon admiration...

Calculer la plus grande communauté forte ? C'est :

- pas vu en Licence
- résoudre le problème de la clique maximum (graphe non orienté)
- un problème conjecturé de complexité exponentielle...

- elle gagnera tout de suite 1 million de dollars (prix CMI)
- et ses membres :
 - un Turing Award, une médaille Fields (s'ils ont moins de 40 ans), un prix Nobel d'Économie, le prix Gödel...
 - un poste de professeur au MIT, ou à Standord, Princeton, Harvard, Oxford, Cambridge, etc... ou éventuellement à AMU
 - et mon admiration...
- elle devra peut-être (sûrement plutôt...) veiller à sa sécutité (si ses membres sont toujours vivants...)

Objectifs très concrets de ce cours :

• Identifier ce qui est "facile" : problèmes polynomiaux

Objectifs très concrets de ce cours :

- Identifier ce qui est "facile" : problèmes polynomiaux
- et ce qui ne l'est peut-être pas : problèmes non-polynomiaux (ou "conjecturés" comme tels)
 et dans ce cas, essayer d'en évaluer la difficulté théorique et pratique

Objectifs très concrets de ce cours :

- Identifier ce qui est "facile" : problèmes polynomiaux
- et ce qui ne l'est peut-être pas : problèmes non-polynomiaux (ou "conjecturés" comme tels)
 et dans ce cas, essayer d'en évaluer la difficulté théorique et pratique

Motivations pour cette UE (un peu théorique) :

- le contenu est nécessaire à tout informaticien (à Bac+5)
- elle vaut 6 crédits!

Objectifs très concrets de ce cours :

- Identifier ce qui est "facile" : problèmes polynomiaux
- et ce qui ne l'est peut-être pas : problèmes non-polynomiaux (ou "conjecturés" comme tels)
 et dans ce cas, essayer d'en évaluer la difficulté théorique et pratique

Motivations pour cette UE (un peu théorique) :

- le contenu est nécessaire à tout informaticien (à Bac+5)
- elle vaut 6 crédits!

et surtout, elle peut vous rapporter 1 million de \$:

Objectifs très concrets de ce cours :

- Identifier ce qui est "facile" : problèmes polynomiaux
- et ce qui ne l'est peut-être pas : problèmes non-polynomiaux (ou "conjecturés" comme tels)
 et dans ce cas, essayer d'en évaluer la difficulté théorique et pratique

Motivations pour cette UE (un peu théorique) :

- le contenu est nécessaire à tout informaticien (à Bac+5)
- elle vaut 6 crédits!

et surtout, elle peut vous rapporter 1 million de \$:

En résolvant la conjecture P ≠ NP
 1 des 7 problèmes des Mathématiques du 21ème siècle

Objectifs très concrets de ce cours :

- Identifier ce qui est "facile" : problèmes polynomiaux
- et ce qui ne l'est peut-être pas : problèmes non-polynomiaux (ou "conjecturés" comme tels)
 et dans ce cas, essayer d'en évaluer la difficulté théorique et pratique

Motivations pour cette UE (un peu théorique) :

- le contenu est nécessaire à tout informaticien (à Bac+5)
- elle vaut 6 crédits!

et surtout, elle peut vous rapporter 1 million de \$:

- En résolvant la conjecture P ≠ NP
 1 des 7 problèmes des Mathématiques du 21ème siècle
- Classe P : problèmes Polynomiaux

Objectifs très concrets de ce cours :

- Identifier ce qui est "facile" : problèmes polynomiaux
- et ce qui ne l'est peut-être pas : problèmes non-polynomiaux (ou "conjecturés" comme tels)
 et dans ce cas, essayer d'en évaluer la difficulté théorique et pratique

Motivations pour cette UE (un peu théorique) :

- le contenu est nécessaire à tout informaticien (à Bac+5)
- elle vaut 6 crédits!

et surtout, elle peut vous rapporter 1 million de \$:

- En résolvant la conjecture P ≠ NP
 1 des 7 problèmes des Mathématiques du 21ème siècle
- Classe P : problèmes Polynomiaux
- Classe NP : problèmes Polynomiaux sur modèle Non-déterministe (NP ne veut pas dire Non-Polynomial !!!)





Bâtiment du CSAIL au MIT - Cambridge (Boston)

Computer Sciences and Arificial Intelligence Laboratory

Coût: 500 millions \$ - Construction en 2003 - Architecte Franck Gerry

Avec un zoom sur la fenêtre en bas à droite

Avec un zoom sur la fenêtre en bas à droite



Le bâtiment avait déjà des fuites d'eau en 2006...

... mais la conjecture $P \neq NP$ tient toujours!

Plan

- Organisation de l'UE
- Évaluation
- 3 Sur la matière enseignée
- Plan du cours

Plan du cours :

Plan du cours :

- Chapitre 1 : Complexité, Présentation et Introduction (c'est presque terminé!)
- Chapitre 2 : Rappels d'algorithmique et de complexité
- Chapitre 3 : Problèmes et Complexité des problèmes
- Chapitre 4 : Cadre formel
- Chapitre 5 : La Classe P (problèmes polynomiaux)
- ... je passe le relai au collègue... Chapitre 6 : La Classe NP
- ...

