université de BORDEAUX	**Année Universitaire 2019/2020, Session 1 d'automne** Master CSI, UE 4TCY701U, Théorie de la complexité. Lundi 16 décembre 2019. 14h30–17h30 Durée : 3h. **Documents non autorisés.**	Collège Sciences et Technologies

Toute réponse non justifiée n'apportera pas de point. Le barème est indicatif.
Vous pouvez utiliser une question précédente même si elle n'a pas été traitée.

Exercice 1 : Décidable vs. indécidable (questions d'application du cours) — 4 Points

Répondez aux questions ci-dessous en **justifiant brièvement mais précisément** (seules les réponses correctement justifiées apportent un point) :

1) Le problème « **Inclusion** » suivant est-il décidable ?
Entrée : Le code de deux machines de Turing M_1 et M_2.
Question : Est-il vrai que $L(M_1) \subseteq L(M_2)$?

2) Le problème suivant est-il dans la classe de complexité **L** ?
Entrée : Pas d'entrée.
Question : Soient $u_1 = 100$, $u_2 = 0$, $u_3 = 1$ et $v_1 = 1$, $v_2 = 100$, $v_3 = 0$. Existe-t-il une suite d'indices $i_1, i_2, \ldots, i_k$ (où $k \geqslant 1$), tous entre 1 et 3, tels que $u_{i_1} u_{i_2} \cdots u_{i_k} = v_{i_1} v_{i_2} \cdots v_{i_k}$?

3) Le problème « **Arrêt universel en k étapes** » suivant est-il décidable ?
Donnée : Une machine de Turing M et un entier k.
Question : La machine M s'arrête-t-elle sur toute entrée en au plus k étapes ?

4) Le problème « **Arrêt universel borné** » suivant est-il décidable ?
Donnée : Une machine de Turing M.
Question : Existe-t-il un entier k tel que la machine M s'arrête sur toute entrée en au plus k étapes ?

Exercice 2 : P vs. NP-complet — 4 Points

On suppose que **P** $\neq$ **NP**. Pour chacun des problèmes suivants, dites si il est dans **P** ou **NP**-complet.

Attention : les réponses doivent être **prouvées**. Si vous répondez qu'un problème est dans **P**, vous devez décrire en français un algorithme polynomial qui le résout et expliquer pourquoi il est polynomial. Si vous répondez qu'un problème est **NP**-complet, vous devez prouver qu'il est dans **NP** *et* qu'il est **NP**-difficile. Pour cela, vous pouvez utiliser n'importe quel problème **NP**-complet parmi ceux vus en cours ou présents dans une feuille de TD, et uniquement ceux-ci.

Remarque *: les réductions sont faciles, la principale difficulté est de trouver le bon problème à réduire.*

1) **SAT modifié**
Entrée : Une formule propositionnelle en forme CNF.
Question : Existe-t-il une affectation des variables qui rend la formule vraie et une autre affectation des variables qui rend la formule fausse ?

2) **Double Hamiltonien**
Entrée : Un graphe non orienté G.
Question : Existe-t-il un chemin dans G qui passe exactement deux fois par chaque sommet ?

3) **Somme approchée**
Entrée : Des entiers strictement positifs $x_1, \ldots, x_n$ et un entier S.
Question : Existe-t-il un sous-ensemble I de $\{1, \ldots, n\}$ tel que $\left| \sum_{i \in I} x_i - S \right| \leqslant 42$?

Exercice 3 : Graphes Eulériens — 4 Points

Dans cet exercice, on considère des graphes ***orientés***. On dit qu'un graphe orienté G est ***fortement connexe*** si pour tous sommets s, t de G, il existe un chemin de s à t dans G. On considère le problème suivant :

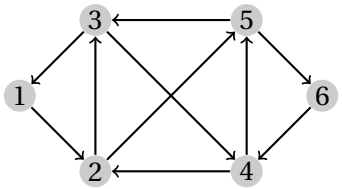
FORTE CONNEXITÉ
Entrée : Un graphe orienté G.
Question : Le graphe G est-il fortement connexe?

1) On sait que l'une des trois propriétés suivantes est vraie. Déterminez laquelle, et prouvez-la.

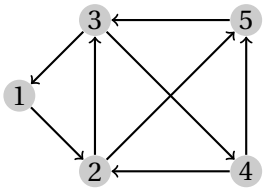
- le problème **FORTE CONNEXITÉ** est **NL**-complet,
- le problème **FORTE CONNEXITÉ** est **P**-complet,
- le problème **FORTE CONNEXITÉ** est **NP**-complet.

2) Est-il possible que ces trois propriétés soient vraies en même temps?

Un ***cycle Eulérien*** dans un graphe orienté G est un cycle passant une et une seule fois par chaque arête (par contre, un tel cycle peut passer plusieurs fois par un même sommet). Le graphe de la figure (a) ci-dessous a un cycle Eulérien : $1 \to 2 \to 3 \to 4 \to 2 \to 5 \to 6 \to 4 \to 5 \to 3 \to 1$. On peut vérifier que le graphe de la figure (b) n'a pas de cycle Eulérien.



(a)



(b)

Le ***degré entrant*** d'un sommet s est le nombre d'arêtes de la forme $t \to s$. Le ***degré sortant*** de s est le nombre d'arêtes de la forme $s \to t$. Le ***degré*** de s est la somme de son degré entrant et de son degré sortant.

3) Montrer qu'un graphe orienté a un cycle Eulérien Eulérien si et seulement si il vérifie les deux conditions suivantes :

- le degré entrant de chaque sommet est égal à son degré sortant, et
- le graphe obtenu en supprimant les sommets de degré 0 est fortement connexe.

4) Donnez la plus petite classe de complexité parmi **L**, **NL**, **P**, **NP**, **PSPACE** contenant le problème suivant :
Entrée : Un graphe orienté G.
Question : Le graphe G a-t-il un cycle Eulérien?

5) Même question si on suppose que le graphe donné en entrée est ***fortement connexe***.

Exercice 4 : Logiques sur les entiers naturels — 10 Points

On considère des formules exprimant des énoncés arithmétiques portant sur les entiers naturels. Un tel énoncé peut être vrai ou faux. Voici trois exemples de formules φ_1, φ_2 et φ_3 :

$$\begin{aligned}
\varphi_1 &= \forall x \exists y \quad (y = x+1).\\
\varphi_2 &= \exists x \exists y \exists z \quad (x^2+y^2=z^2 \wedge \neg(x\,y\,z=0)).\\
\varphi_3 &= \exists x \exists y \exists z \quad (x^3+y^3=z^3 \wedge \neg(x\,y\,z=0)).
\end{aligned}$$

Dans ces exemples, on a quantifié des variables (ici, x, y, z) existentiellement ($\exists$) ou universellement ($\forall$). Les variables sont interprétées comme des ***entiers naturels***. Ces exemples de formules utilisent les constantes 0 et 1, l'égalité, l'addition, la multiplication (présente dans les deux dernières formules), ainsi que des connecteurs logiques (ici, $\neg$ et $\wedge$).

La formule φ_1 énonce que tout entier naturel a un successeur : elle est vraie. La formule φ_2 est également vraie, comme le montre le choix $x=3$, $y=4$ et $z=5$ (on a bien $3^2+4^2=5^2$ et $3\times4\times5\neq0$). Enfin, la formule φ_3 est fausse (c'est un cas très particulier du grand théorème de Fermat qu'on peut montrer directement).

Certaines variables peuvent aussi être ***libres***, c'est-à-dire non quantifiées. Par exemple, la formule,

$$\varphi_4(x) \quad = \quad \exists y \quad (x = y + y)$$

a une variable libre x, ce qu'on indique explicitement en la notant $\varphi_4(x)$. L'autre variable y est quantifiée : on dit qu'elle est ***liée***. La valeur de φ_4 dépend de celle de x : si x pair, la formule est vraie (car dans ce cas, il existe un entier naturel y tel que $x = y + y$). Au contraire, si x est impair, $\varphi_4(x)$ est fausse. La formule φ_4 est donc vraie si et seulement si x est pair : elle exprime ainsi la propriété « *l'entier naturel x est pair* ». Plus généralement, lorsqu'une formule φ a des variables libres $y_1, \ldots, y_k$, on la note aussi $\varphi(y_1, \ldots, y_k)$, et sa valeur de vérité dépend des entiers naturels $y_1, \ldots, y_k$: elle définit une propriété sur les k-uplets d'entiers.

Cet exercice considère plusieurs syntaxes permises pour les formules. La première est la suivante : une formule est de la forme,

$$Q_1 x_1 Q_2 x_2 \cdots Q_n x_n \, \varphi,$$

où chaque Q_i est soit $\exists$, soit $\forall$, et où φ est une formule sans quantificateur, construite à partir de formules « *de base* » en utilisant les connecteurs logiques $(\wedge, \vee, \neg, \Rightarrow, \Leftrightarrow)$. Les formules de base sont de la forme $t = t'$, où t et t' sont appelés des « ***termes*** ». Chaque terme est une somme de variables. Par exemple, $(x = x + x + x)$ et $(x + y = z + z)$ sont des formules de base (x, y, z sont des variables). Une variable de φ peut être,

— soit l'une des variables quantifiées $x_1, \ldots, x_n$: on dit qu'une telle variable est ***liée***.

— soit une autre variable : dans ce cas, elle n'est liée a aucun quantificateur Q_i, on dit qu'elle est ***libre***.

Note. La formule φ_4 est conforme à cette syntaxe, mais pas la formule φ_1 (qui utilise la constante 1, pour l'instant interdite) ni les formules φ_2 et φ_3 (qui utilisent la constante 0 et des produits de variables). Nous allons d'abord montrer qu'on peut en fait s'autoriser les constantes.

Partie I : Exemples de propriétés.

1) Soit k un entier naturel. Définir par récurrence sur k une formule $\alpha_k(x)$, avec une variable libre x, telle que $\alpha_k(x)$ est vraie si et seulement si $x = k$.

La question précédente permet de généraliser la syntaxe des termes à des combinaisons affines de variables, c'est-à-dire de la forme $k + \sum_{i=1}^{p} k_i x_i$ où $k, k_i \in \mathbb{N}$ (sans pour autant augmenter l'ensemble des propriétés mathématiques exprimables). Par exemple, la formule $(x = 3y + 2)$ s'écrit, avec les formules de base précédentes, comme $\exists z(\alpha_2(z) \wedge (x = y + y + y + z))$. On a utilisé une variable z et α_2 pour définir la constante 2. On permet donc les combinaisons affines de variables à partir de maintenant.

2) Écrire une formule à deux variables libres x, y qui exprime la propriété $x \leqslant y$. Vous pouvez utiliser une ou plusieurs autres variables (liées). De même, écrire une formule qui exprime la propriété $x < y$.

La question **2)** permet désormais d'utiliser les comparaisons de termes $t < t'$, $t \leqslant t'$, $t > t'$ et $t \geqslant t'$.

3) Exprimez en français ce qu'énonce la formule $\forall x \exists y \exists z \, (y > x) \wedge (y = 2z + 1)$. Est-elle vraie ?

> **Attention !**
> On ne demande pas une paraphrase de cette formule (du type « pour tout entier x, il existe un entier y et un entier z tels que... ») mais d'expliquer quelle propriété des entiers naturels est exprimée.

Dans les deux questions suivantes, on autorise les formules de base de la forme $x = yz$, où x, y, z sont des variables, en plus des formules de base précédentes.

4) Quelle est la propriété sur z exprimée par la formule $\kappa(z) = \exists x \exists y \, (x > 1) \wedge (y > 1) \wedge (z = xy)$?

5) Écrire une formule exprimant qu'il existe une infinité de couples $(p, p+2)$ où p et $p+2$ sont premiers.

Dans le reste de l'exercice, on s'intéresse au problème de calculer la valeur de vérité de telles formules sans variable libre (une telle formule est donc vraie ou fausse), du point de vue décidabilité et complexité. Dans les parties II et III, on considère ce problème en restreignant, de façon différente dans chaque partie, l'ensemble des formules d'entrée par la syntaxe autorisée.

Partie II : Logique avec égalité seulement.

On considère maintenant des formules dont la syntaxe est restreinte : les seuls termes autorisés sont les constantes (codées en binaire) et les variables. Ainsi, $x = y$, $x = 0$, $x = 1$ et $x = 42$ sont des formules de base.

1) Quelle est la complexité la plus précise que vous connaissez contenant le problème suivant ?

Entrée : Une formule φ dont la syntaxe est restreinte comme ci-dessus, sans variable libre.
Question : La formule φ est-elle vraie ?

2) Même question si on suppose en plus que chaque quantificateur de la formule d'entrée est $\exists$.

On revient maintenant à la « *logique avec addition* », dans laquelle les termes autorisés sont des combinaisons affines de variables.

3) Montrer qu'il est **PSPACE**-difficile de déterminer si une telle formule, donnée en entrée, est vraie.

Partie III : Logique avec addition et multiplication.

Dans cette partie, on étend la syntaxe de la façon suivante : un terme est un ***produit*** d'une ou plusieurs combinaisons affines de variables à coefficients entiers naturels. Par exemple, $5z + 1 = (2x + 1)(3y + 2)$ est une formule de base autorisée. On appelle cette logique « *logique avec addition et multiplication* ». Les formules φ_1, φ_2, φ_3 et φ_4 sont conformes à cette syntaxe. L'objectif est de montrer que le problème « LOGIQUE ADD-MULT » suivant est indécidable :

Entrée : Une formule φ sans variable libre dans la logique avec addition et multiplication.
Question : La formule φ est-elle vraie ?

À tout mot w sur l'alphabet $\{1,2\}$, on associe sa valeur $\overline{w}$ en base dix. Par convention, la valeur du mot vide est $\overline{\varepsilon} = 0$. Ainsi par exemple, $\overline{121}$ vaut l'entier 121 (cent vingt et un).

1) Que peut-on dire de u et v lorsque $\overline{u} = \overline{v}$?

2) Soit u, w deux mots de $\{1,2\}^*$. On pose $w = a_1 \cdots a_n$. Exprimer $\overline{uw}$ en fonction de $\overline{u}$, $\overline{w}$ et de n.

3) Soit $w \in \{1,2\}^*$. Écrire une formule $f_w(x,y)$ à deux variables libres x, y qui exprime que l'entier y est la valeur du mot obtenu en concaténant w à la suite de la représentation décimale de x. Par exemple, si w est le mot 2222, $f_w(x,y)$ est vraie pour les entiers $x = 121$ et $y = 1212222$ (et, pour cette valeur de x, $f_w(x,y)$ n'est vraie pour aucune autre valeur de y).

4) On suppose dans cette question seulement que l'on peut quantifier l'existence d'une suite finie d'entiers naturels. Autrement dit, on peut écrire $\exists n \, \exists (m_k)_{k \leqslant n} \varphi$. Montrer que le problème suivant est indécidable :

Entrée : Une formule φ sans variable libre dans la logique avec addition, multiplication, et quantification sur les suites.
Question : La formule φ est-elle vraie ?

Indication

Utiliser les questions **1)** et **3)** pour réduire le problème de correspondance de Post à ce problème.

On veut maintenant se passer de la quantification sur les suites finies pour montrer l'indécidabilité du problème LOGIQUE ADD-MULT. On utilise le théorème des restes chinois dont l'énoncé est le suivant.

> Soient $n_1, \ldots, n_k$, $x_1, \ldots, x_k$ des entiers naturels. On suppose que si $i \neq j$, n_i et n_j sont premiers entre eux. Alors, il existe un entier x tel que $x \equiv x_i \pmod{n_i}$ pour tout $1 \leqslant i \leqslant k$.

Soit $(x_1, \ldots, x_k)$ une suite finie d'entiers naturels et $n = (\max(k, x_1, \ldots, x_k))!$, où $p!$ désigne la factorielle de p. Pour chaque $i = 1, \ldots, k$, soit $n_i = in + 1$.

5) Montrer que si $i \neq j$, les entiers n_i et n_j sont premiers entre eux.

6) Soit x un entier donné par le théorème des restes chinois. Montrer que chaque x_i peut s'exprimer en fonction de x, n, i en utilisant uniquement l'addition et la multiplication.

7) Déduire des questions précédentes que le problème LOGIQUE ADD-MULT est indécidable.