Fondements de l'informatique. Examen
Durée: 3h

Sujet proposé par Olivier Bournez
Version 5

L'énoncé comporte 5 parties (sections), certaines avec des sous-parties (sous-sections), chacune indépendante, qui pourront être traitées dans un ordre quelconque. En revanche, dans chaque partie, il peut être utile, dans la réponse à une question, d'utiliser les questions précédentes ! On pourra librement admettre le résultat d'une question pour passer aux questions suivantes. La difficulté des questions n'est pas une fonction linéaire ni croissante de leur numérotation.

La qualité de votre argumentation et de votre rédaction est une partie importante de votre évaluation.

On pourra utiliser les résultats et les théorèmes démontrés en cours sans chercher à les redémontrer.

Dans tout l'énoncé, on demande des algorithmes et des solutions à un haut niveau : dans aucune des questions il n'est demandé de décrire complètement une machine de Turing, ni même d'en donner une description graphique ; on pourra se contenter pour décrire un algorithme de le décrire par exemple en français ou dans un langage de programmation classique comme PYTHON, JAVA, C ou OCAML.

1 Calculabilité

Dans cette section, on considère que l'alphabet est $\{A, B, \ldots, Z, 0, \ldots, 9\}$.

Question 1. *Peut-on décider si une machine de Turing accepte le mot* INF412 *? Justifier.*

Question 2. *Peut-on décider si une machine de Turing accepte le mot* INF412 *en moins de 412 étapes ? Justifier.*

Question 3. *Soit A un langage récursivement énumérable. On suppose que A contient exactement 412 mots de taille n pour tout n. A est-il décidable ? Justifier.*

Question 4. *Soit A un langage indécidable. Le langage $B = \{n | \exists x \in A, |x| \geq n\}$, où n est un entier codé en binaire, et $|x|$ désigne la longueur du mot x, est-il décidable ? Justifier.*

2 Ordres

On considère une signature Σ qui contient un symbole de relation binaire $\leq$, et un symbole binaire d'égalité $=$.

Question 5. *Ecrire des formules sur la signature Σ qui expriment que la relation $\leq$ est une relation d'ordre (i.e. est réflexive, transitive et antisymétrique).*

Un ordre est total si pour tout x et y, on a $x \leq y$ ou $y \leq x$.

Tout ensemble *fini* peut être muni d'une relation d'ordre total : chaque énumération d'un ensemble fini détermine un ordre total sur cet ensemble. Il y a donc $n!$ ordres total pour un ensemble fini de n éléments.

Question 6. *Utiliser le théorème de compacité pour démontrer que tout ensemble peut-être muni d'une relation d'ordre total.*

3 Principe de Robinson

On rappelle qu'un *corps commutatif* est un modèle des axiomes (6.8) à (6.17) du polycopié. On rappelle qu'un corps de *caractéristique* p est un corps où $\mathbf{1} + \mathbf{1} + \cdots + \mathbf{1} = \mathbf{0}$ où le $\mathbf{1}$ est répété p fois, et p est le plus petit entier avec cette propriété. Un corps est dit de *caractéristique* 0 s'il n'est pas de caractéristique p pour tout entier p. Par exemple, on connaît $\mathbb{Q}$ qui est de caractéristique 0, et $\mathbb{Z}/p\mathbb{Z}$ de caractéristique p.

Question 7. *Démontrer que si une formule ϕ du premier ordre*[1] *est vraie dans tous les corps de caractéristique 0, alors il existe un entier k tel que la formule ϕ est vraie dans tous les corps de caractéristique supérieure ou égale à k.*

4 NP-complétude

L'objectif de cet exercice est de prouver que le problème de décision SC suivant est NP-complet :

- **Donnée**: Une formule F' du calcul propositionnel sous la forme d'une conjonction de clauses de 2 ou 3 littéraux, avec **chaque variable qui apparaît au total au plus trois fois** (en comptant ses occurences positives et négatives) .
- **Réponse**: Décider si F' est satisfiable.

Question 8. *Soit x_1, $x_2 \dots x_k$ des variables propositionnelles. Considérons la formule ϕ définie par la conjonction des k clauses $(x_i \vee \neg x_{i+1})$ pour $i = 1, \dots, k-1$, et de la clause $(x_k \vee \neg x_1)$. Quelles sont les assignations des variables qui satisfont ϕ ?*

Question 9. *Prouver que le problème SC est NP-complet.*

5 Théorème de Fagin

L'objectif de cette section est de prouver qu'il est possible de définir NP sans aucune notion de machine, par la logique du second ordre existentiel. Ce résultat, dû à Fagin [1] est fondateur de la complexité descriptive.

5.1 Quelques concepts vus par un logicien

La logique du premier ordre (le calcul des prédicats) vu en cours permet d'exprimer certains concepts facilement.

Par exemple, pour le logicien, un **graphe (non orienté et sans boucle)** se définit comme un modèle sur une signature Σ_G contenant un symbole E de relation d'arité 2, qui satisfait la formule

$$\forall x \forall y((\neg E(x,x)) \wedge (E(x,y) \Rightarrow E(y,x))). \tag{1}$$

Autre exemple, puisqu'une forme normale conjonctive est au final une conjonction de clauses, chaque clause étant une disjonction de littéraux, pour le logicien, une **forme normale conjonctive** se définit comme un modèle sur une signature Σ_{FNC} contenant les symboles de relation P

1. C'est-à-dire une formule du calcul des prédicats vu en cours.

et N d'arité 2 : son domaine est vu comme un ensemble de clauses et de variables et la relation $P(c, v)$ signifie que la variable v apparaît positivement dans la clause c, et $N(c, v)$ signifie que la variable v apparaît négativement dans la clause c.

Question 10. *Décrire complétement la (une) structure qui correspond à la forme normale conjonctive*

$$(x_1 \vee \neg x_2 \vee x_3) \wedge (\neg x_1 \vee \neg x_2 \vee x_4) \wedge (x_2 \vee \neg x_3 \vee x_5).$$

5.2 Logique du second ordre

La logique du second ordre consiste à ajouter à la logique du premier ordre des variables de relations sur lesquelles on peut quantifier : Le principe est que la formule $\exists A^r\ \phi$ signifie qu'il y a un choix de relation A d'arité r, tel que la formule ϕ est satisfaite.

On va se limiter dans cet énoncé aux formules $SO\exists$: c'est-à-dire avec des quantifications existentielles sur des symboles de relations, placées au début. [2]

Détails formels si nécessaire : Fixons une signature [3] Σ.

— Une formule de $SO\exists$ est de la forme

$$\exists R_1^{n_1}\ \exists R_2^{n_2}\ \ldots \exists R_k^{n_k}\ \psi$$

où ψ est une formule du premier ordre sur la signature $\Sigma \cup \{R_1^{n_1}, \ldots, R_k^{n_k}\}$, où $\Sigma \cup \{R_1^{n_1}, \ldots, R_k^{n_k}\}$ est la signature Σ à laquelle on a ajouté les symboles de relations R_1, R_2, ..., R_k d'arités respectives n_1, n_2, ..., n_k.

— Une telle formule est satisfaite en une structure sur la signature Σ s'il existe un choix d'interprétation pour les relations R_1, R_2, ..., R_k avec les arités correspondantes telle que la formule ψ est satisfaite.

Prenons l'exemple suivant (comme ci-dessus, les exposants dans les quantifications sur les symboles de relation indiquent l'arité du symbole).

$$\exists B^1 \exists N^1 \exists R^1 \forall x((B(x) \vee N(x) \vee R(x)) \wedge (\forall y(E(x, y) \Rightarrow$$
$$(\neg(B(x) \wedge B(y)) \wedge \neg(N(x) \wedge N(y)) \wedge \neg(R(x) \wedge R(y))))))$$

Question 11. *Cette formule $\phi_{3-COLORABILITE}$ de $SO\exists$ exprime le problème $3-COLORABILITE$: Expliquer pourquoi un graphe est coloriable avec 3-couleurs si et seulement, lorsqu'on le voit comme une structure sur Σ_G, [4] cette formule $\phi_{3-COLORABILITE}$ est satisfaite en cette structure ? à quoi correspondent B, N et R ?*

Question 12. *Donner une formule ϕ_{SAT} de $SO\exists$ qui exprime le problème SAT : c'est-à-dire telle qu'une forme normale conjonctive est satisfiable si et seulement si, lorsqu'elle est vue comme une structure sur Σ_{FNC}, [5] cette formule ϕ_{SAT} est satisfaite en cette structure.*

On se focalise dans toute la suite sur les modèles égalitaires finis et ordonnés, c'est-à-dire dont l'ensemble de base [6] est fini, et avec un symbole $=$ d'arité 2, interprété par l'égalité, et un symbole $\leq$ d'arité 2 avec une interprétation satisfaisant les axiomes habituels d'ordre total. Pour simplifier, on ne considérera que des signatures sans symboles de fonctions. [7]

2. Voir note bibliographique pour le cadre général de la logique du second ordre.
3. k, n_1, ..., n_k sont bien entendu des entiers.
4. Voir la section 5.1.
5. Voir la section 5.1.
6. synonyme domaine.
7. Les signatures ne possèdent donc que des symboles de constantes et des symboles de relations.

Un graphe *ordonné* (non orienté et sans boucle) est un graphe au sens précédent dont les sommets sont ordonnés : c'est-à-dire un modèle sur une signature $\Sigma_G \cup \{=, \leq\}$, où $\Sigma_G \cup \{=, \leq\}$ désigne la signature Σ_G à laquelle sont ajoutés des symboles de relations $=$ et $\leq$ interprétés par l'égalité et une interprétation satisfaisant les axiomes habituels d'ordre total, et qui satisfait la formule (1).

Question 13. *Soit $\Sigma_G \cup \{=, \leq, s\}$ la signature $\Sigma_G \cup \{=, \leq\}$ avec une nouvelle constante s. Une structure finie sur cette signature comme dans le paragraphe précédent est donc la donnée (G, a) d'un graphe G fini ordonné et de l'un de ses sommets a interprétant s.*

Donnner une formule du premier ordre sur cette signature[8] *telle que : pour tout graphe G et tout sommet a de G, les sommets plus petits que a forment une clique*[9] *si et seulement si la formule est satisfaite en (G, a).*

Question 14. *Donner une formule ϕ_{CLIQUE} de $SO\exists$ qui exprime le problème $CLIQUE$: c'est-à-dire une formule sur la signature $\Sigma_G \cup \{=, \leq, s\}$ telle que pour tout graphe ordonné fini G et pour tout entier k, G possède une clique de taille k si et seulement si ϕ_{CLIQUE} est satisfaite en la structure (G, a), où a est l'interprétation de s comme le $k$ième sommet*[10] *de G.*

(Indication : on pourra pour cela identifier et utiliser une notion de relation bijective)

Tous ces problèmes sont des problèmes de NP.

Le théorème de Fagin consiste à observer que les problèmes de décision qui s'expriment par une formule de $SO\exists$ de cette façon sont exactement les problèmes de NP.

5.3 Démonstration du théorème de Fagin

Pour formaliser cela, il faut arriver à parler de problèmes de décision sur des structures, et donc fixer une représentation.

Fixons une signature Σ.

Soit $\mathcal{M}$ une structure finie ordonnée sur cette signature. Puisque les éléments du domaine sont en nombre fini et ordonnés, on peut les appeler $1, 2, \ldots, n$ pour un certain n. Chaque symbole de relation $R \in \mathcal{R}$ d'arité k s'interprète comme un sous-ensemble de $\{1, 2, \ldots, n\}^k$. On peut le coder par un mot sur l'alphabet $\{0, 1\}$ de longueur n^k, où le 1 à la $i$ème position indique que le k-uplet correspondant est dans le sous-ensemble. De même chaque symbole de constante peut se coder par le codage en binaire de son interprétation. Le codage de $\mathcal{M}$, noté $< \mathcal{M} >$, est alors la concaténation des codages de ses constantes et relations.[11]

Question 15. *Soit ϕ une formule du premier ordre sur une signature Σ. Démontrer qu'il y a un algorithme qui prend en entrée le codage $< \mathcal{M} >$ d'une structure finie ordonnée sur Σ, et qui décide en temps polynomial si $\mathcal{M}$ satisfait ϕ.*

On va tout d'abord démontrer le sens le plus facile du théorème de Fagin : toute problème de décision qui s'exprime par une formule de $SO\exists$ est dans NP.

Question 16. *Formellement : soit Σ une signature, et soit ϕ une formule de $SO\exists$. Démontrer que le problème suivant est dans NP :*

- ***Donnée**: le codage $< \mathcal{M} >$ d'une structure finie ordonnée sur Σ.*
- ***Réponse**: décider si $\mathcal{M}$ satisfait la formule ϕ.*

Question 17. *Que peut-on dire de plus sur la difficulté du problème précédent ?*

8. C'est-à-dire sur la signature $\Sigma_G \cup \{=, \leq, s\}$.
9. On rappelle qu'une clique est un sous-ensemble de sommets reliés deux-à-deux.
10. Puisque les sommets de G sont ordonnés, et en nombre fini, on peut les numéroter de 1 à n. Le $k$ième est donc celui de numéro k dans l'ordre total donné par $\leq$.
11. Rappel : on considère des signatures sans symboles de fonctions pour simplifier.

On va maintenant montrer l'autre direction : tout problème A de NP s'exprime par une formule de $SO\exists$.

Formellement : soit Σ une signature. Soit A un problème de décision sur les structures finies et ordonnées sur Σ qui est dans NP. Par définition, cela veut dire que l'on a $< \mathcal{M} > \in A$ si et seulement si $\exists u$, V accepte (w, u), pour un certain vérificateur polynomial V, où $w =< \mathcal{M} >$.

Question 18. *Construire une formule ϕ_A de $SO\exists$ qui exprime que V possède un calcul accepteur sur $w\#u$ pour un certain u, où $w =< \mathcal{M} >$.*

(on pourra admettre dans un premier temps le fait qu'il est possible de définir un ordre[12] *sur les k-uplets à partir de l'ordre $\leq$ sur les éléments, et on pourra (alors) noter $t+1$ pour le successeur de t).*

La direction manquante du théorème de Fagin en découle : Soit A un problème de décision sur les structures finies et ordonnées sur Σ qui est dans NP. La formule ϕ_A exprime le problème $A : \mathcal{M}$ est une instance positive de A si et seulement s'il satisfait la formule ϕ_A.

Notes bibliographiques

Sur la logique du second ordre

On s'est restreint dans cette énoncé aux quantifications du second ordre existentielles. Mais en logique du second ordre, on autorise aussi en général les quantifications universelles. Par exemple, la formule $\forall A^r\ \phi$ signifie que pour tout choix de relation A d'arité r, la formule ϕ est satisfaite.

Toute formule du second ordre peut se transformer en une formule équivalente où toutes les quantifications du second ordre sont au début.

Sur le théorème de Fagin

Le théorème de Fagin est un résultat remarquable, car on voit qu'il permet de définir NP sans avoir à définir la moindre notion de machine (par exemple de machine de Turing), ou d'algorithme ! C'est assez inattendu ! Il a donné naisance à ce que l'on appelle la complexité descriptive et complexité implicite, où des définitions sans notion de machine des principales classes de complexité ont été obtenues (P, $PSPACE$, ...).

Sur la partie 5

La partie 5 est inspirée de [2].

Références

[1] R. Fagin. Generalized first-order spectra and polynomial-time recognizable sets. In R. M. Karp, editor, *Complexity in Computer Computations*, pages 43–73. American Mathematics Society, Providence R.I., 1974.

[2] N. Immerman. *Descriptive Complexity.* Springer, 1999.

12. Par exemple l'ordre lexicographique