Fondements de l'informatique. Examen Durée: 3h

Sujet proposé par Olivier Bournez
Version 7

Les 5 parties sont indépendantes, et peuvent être traitées dans un ordre quelconque. On pourra admettre le résultat d'une question pour passer aux questions suivantes. On pourra utiliser tous les résultats et les théorèmes démontrés ou énoncés en cours ou en petite classe ou dans le polycopié sans chercher à les redémontrer, sauf si cela est explicitement demandé.

Il est possible d'avoir la note maximale sans répondre à toutes les questions. La difficulté des questions n'est pas une fonction linéaire ni croissante de leur numérotation.

La ***qualité et clarté de votre argumentation et de votre rédaction*** *sera une partie importante de votre évaluation.*

Il peut être clair pour vous que certaines notions impliquent d'autres notions, mais à une question du type "le problème est-il A ? est-il B ? " ***il est demandé de répondre aux deux questions****, soit en répondant à chacune, soit en rendant explicite que vous savez que A implique B ou le contraire pour avoir la note maximale, possiblement une fois pour toute dans votre copie.*

On note $|S|$ pour le nombre d'éléments de l'ensemble (fini) S.

1 Décidabilité

On considère des machines de Turing qui travaillent sur un alphabet suffisant pour couvrir tous les symboles utilisés dans un traitement de texte.

Question 1. *Parmi les problèmes de décision suivants, lesquelles sont décidables ? Récursivement énumérable ? Dans NP, NP-complet, dans P ? Justifier votre réponse.*

1. *Déterminer si un sujet d'examen contient le mot "machine de Turing".*
2. *Déterminer si une machine de Turing produit sur l'entrée correspondant au mot vide un sujet d'examen qui contient le mot "machine de Turing".*

2 Logique et décidabilité

Question 2. *On considère le problème $SAT_{propositionnel}$ de décision suivant :*
— ***Donnée****: Une formule F du calcul propositionnel.*
— ***Réponse****: F est-elle satisfiable.*
Ce problème est-il décidable ? Récursivement énumérable ? Dans NP, NP-complet, dans P ? Justifier votre réponse.

Le but est d'étudier la difficulté du problème similaire $SAT_{\text{prédicat}}$ pour la logique du premier ordre [1].

On appelle *machine de Turing à demi-ruban infini*, une machine qui n'écrit jamais aucun symbole à gauche de son entrée : la partie à gauche de l'entrée reste donc constituée à jamais uniquement de symboles de blanc.

1. Rappelons qu'on dit qu'une formule close est satisfiable si et seulement si elle possède un modèle.

Question 3. *Montrer que le problème de décision suivant est indécidable : Etant donnée M une machine de turing à demi-ruban infini, décider si M accepte le mot vide.*

Etant donné un mot $w = a_1 a_2 \dots a_n$ sur l'alphabet Σ, on notera $\overline{w}$ pour le mot w renversé : $\overline{w} = a_n a_{n-1} \dots a_2 a_1$. Pour a un symbole fixé quelconque de Σ, on note a^T pour le mot de longueur T dont toutes les lettres sont le symbole a. En particulier, $a^0 = \epsilon$.

Etant donnée une machine de Turing M, on considère une signature Σ_M constituée du symbole de constante ϵ, d'un symbole de fonction a d'arité 1 pour chaque lettre a de Σ, et d'un symbole f_q d'arité 2 pour chaque état interne $q \in Q$ de M.

Une telle signature Σ_M possède un modèle $\mathfrak{M}_M$ particulier : son ensemble de base est constitué des mots sur l'alphabet Σ (et donc les variables x s'interprètent comme des mots sur l'alphabet Σ), ϵ s'interprète comme le mot vide, $a(w)$ est défini comme le mot aw pour tout mot w, et $f_q(x, y)$ est défini comme vrai si et seulement si M sur l'entrée vide atteint une configuration où elle est dans l'état q, et où le contenu du ruban est $\overline{x}y$ avec la tête de lecture en face de la première lettre de y.

Bien entendu, Σ_M peut posséder d'autres modèles. Mais le théorème de complétude permet d'affirmer qu'une formule F du calcul des prédicats (logique du premier ordre) est prouvable si et seulement si elle est vraie dans tout modèle, c'est-à-dire quelle que soit l'interprétation des symboles de constantes, fonctions et relations.

Question 4. *Soit M une machine de Turing à demi-ruban infini. Montrer que l'on peut construire une formule close F_M qui est prouvable (vraie dans tout modèle) si et seulement si la machine de Turing à demi-ruban infini M accepte le mot vide.*

Question 5. *Montrer qu'une formule close F du calcul des prédicats (logique du premier ordre) est satisfiable si et seulement si sa négation $\neg F$ n'est pas prouvable.*

Question 6. *On considère le problème $SAT_{prédicat}$ de décision suivant :*

- ***Donnée**: Une formule close F du calcul des prédicats (logique du premier ordre).*
- ***Réponse**: F est-elle satisfiable.*

Ce problème est-il décidable ? Récursivement énumérable ? Dans NP, NP-complet, dans P ? Justifier votre réponse.

Question 7. *La signature précédente possède des symboles d'arité 1 et 2.*

Etant donné un entier k, on considère le problème $SAT_{prédicat,k}$ de décision suivant :

- ***Donnée**: Une formule close F du calcul des prédicats (logique du premier ordre) sur une signature avec des symboles d'arité au plus k.*
- ***Réponse**: F est-elle prouvable.*

Montrer que pour certains entier k que l'on précisera, le problème $SAT_{prédicat,k}$ est RE-complet : il est récursivement énumérable, et pour tout problème A récursivement énumérable, $A \leq_m SAT_{prédicat,k}$.

On admettra qu'il est décidable pour $k = 1$.

3 NP-complétude

On s'intéresse dans cette question au problème de décision suivant : Set Packing :

- **Donnée**: Un ensemble fini U, un ensemble $\mathcal{C}$ composé de sous-ensembles $S_1, \dots, S_t$ de U, et un entier z.
- **Réponse**: Existe-il z éléments de $\mathcal{C}$ deux-à-deux disjoints ?

Rappelons qu'une *couverture de sommets* S d'un graphe G est un sous-ensemble de sommets tel que toutes les arêtes de G ont au moins une extrémité dans S.

Partant d'un graphe $G = (V, E)$, nous allons construire un ensemble $\mathcal{C}$ composé de $|V|$ éléments de la façon suivante : pour chaque sommet $v \in V$, on définit l'ensemble S_v comme l'ensemble de toutes les arêtes qui ont v comme extrémité.

Question 8. *Montrer que le graphe possède une couverture de sommets avec k sommets si et seulement si $n - k$ éléments de $\mathcal{C}$ sont deux-à-deux disjoints.*

Question 9. *Montrer que le problème* SET PACKING *est NP-complet.*

On utilisera la NP-complétude du problème Couverture de sommets (VC)

— ***Donnée**: Un graphe non-orienté $G = (V, E)$ et un entier k.*

— ***Réponse**: G admet il une couverture de sommets $\mathcal{S}$ avec au plus k sommets ?*

4 Théorème de Hall : du cas fini au cas infini

On note $\mathbb{N}^+$ pour l'ensemble des entiers naturels non-nuls. On considère $I = \mathbb{N}^+$ ou $I = \{1, 2, \ldots, n\}$ dans tout ce qui suit.

Etant donné un ensemble S et une famille $(S_i)_{i\in I}$ de sous-ensembles de S, un *système de représentants distincts* est un choix d'éléments (c'est-à-dire une famille $(x_i)_{i\in I}$) avec pour tout i, $x_i \in S_i$ tel que pour $i \neq j \in I$, on a $x_i \neq x_j$.

Par exemple, pour $S = \mathbb{N}^+$, $I = \mathbb{N}^+$, et la famille $(S_i)_{i\in\mathbb{N}^+}$, définie par $S_i = \{i, i + 1\}$, on a $S_1 = \{1, 2\}$, $S_2 = \{2, 3\}$, ..., on peut prendre comme système de représentants distincts $x_i = i$.

On admettra le résultat suivant [2], pour le cas où la famille (c'est-à-dire I) est finie :

Théorème [Hall 1935]. Considérons un entier $n > 0$. Une condition nécessaire est suffisante pour l'existence d'un système de représentants distincts pour une famille $(S_i)_{1\leq k\leq n}$ est la suivante :

(H_n) Pour tout $1 \leq k \leq n$, et tout choix d'indices distincts $1 \leq i_1, \ldots, i_k \leq n$, on a $|S_{i_1} \cup \cdots \cup S_{i_k}| \geq k$.

Dans cette partie, on cherche à étendre ce théorème au cas d'une famille infinie.

Question 10. *Une première tentative serait de penser qu'une condition nécessaire et suffisante pour l'existence d'un système de représentants distincts pour une famille $(S_i)_{i\in\mathbb{N}^+}$ est la suivante :*

(H_∞) *Pour tout $k \in \mathbb{N}^+$, et tout choix d'indices distincts $i_1, \ldots, i_k \in \mathbb{N}^+$, on a*

$$|S_{i_1} \cup \cdots \cup S_{i_k}| \geq k.$$

Montrer que cela ne suffit pas, en exhibant une famille de sous-ensembles de $\mathbb{N}$ (les entiers naturels) pour lesquels il n'existe pas de système de représentants distincts. On pourra choisir $S_1 = \mathbb{N}$, et les autres S_i finis.

Par contre, cela fonctionne si l'on suppose que chaque S_i est **finie**. En effet, nous allons prouver le résultat suivant :

Théorème. On considère une famille $(S_i)_{i\in\mathbb{N}^+}$ d'ensembles **finis** de S .

Une condition nécessaire et suffisante pour l'existence d'un système de représentants distincts pour une famille $(S_i)_{i\in\mathbb{N}^+}$ est la suivante :

(H_∞) Pour tout $k \in \mathbb{N}^+$, et tout choix d'indices distincts $i_1, \ldots, i_k \in \mathbb{N}^+$, on a

$$|S_{i_1} \cup \cdots \cup S_{i_k}| \geq k.$$

Question 11. *Prouver que (H_∞) est une condition nécessaire.*

Question 12. *Prouver l'autre direction du théorème. Indication : On pourra chercher à définir une théorie qui est satisfiable si et seulement si la famille admet un système de représentants distincts.*

2. Qui n'est pas difficile à démontrer.

5 Hiérarchie arithmétique : vision logique

Dans cette partie, on considère des formules sur la signature $\mathcal{L} = (0, s, +, \times, <, =)$ de l'arithmétique.

Une formule sur cette signature est dite Σ_0 si elle appartient au plus petit ensemble contenant les formules atomiques et clos par :

— conjonction $\wedge$ (finie), disjonction $\vee$ (finie), négation $\neg$;

— quantification universelle bornée : si φ est Σ_0, si t est un terme, et x une variable n'apparaissant pas dans t, alors $\forall x \leqslant t\ \varphi$ est Σ_0 (ici, $\forall x \ \leqslant t\ \varphi$ est une abréviation pour $\forall x\ (x \leqslant t \Rightarrow \varphi)$) ;

— quantification existentielle bornée, que l'on définit de façon similaire.

Par convention, on considère que Π_0 et Σ_0 sont des synonymes.

Pour $n \in \mathbb{N}$, on dit d'une formule qu'elle est :

— Σ_{n+1} si elle est de la forme $\exists x\ \varphi$ pour φ une formule Π_n ;

— Π_{n+1} si elle est de la forme $\forall x\ \varphi$ pour φ une formule Σ_n.

L'intuition est que Σ (respectivement : Π) signifie que l'on commence par une quantification existentielle (respectivement universelle), et l'indice indique le nombre d'alternances de quantificateurs, sans compter les quantifications bornées. Par exemple, $\exists n \forall m \leq n\ n \leq m \times m$ est une formule Σ_1, et $\exists n \forall j \forall m \leq n\ m \leq j \times n$ est une formule Σ_2.

On cherche à comprendre les sous-ensembles de $\mathbb{N}^p$ que l'on arrive à définir avec ces formules : on dit qu'une formule $\phi(x_1, \ldots, x_p)$ a p-variables libres définit un sous-ensemble de $\mathbb{N}^p$: cet ensemble correspond à l'ensemble des p-uplets qui la satisfait. Un ensemble est alors dit Σ_n (respectivement : Π_n) définissable si ϕ est une formule Σ_n (respectivement : Π_n). On dit qu'il est Δ_n si il est à la fois Σ_n et Π_n définissable.

Pour $n = 0$, on peut se convaincre par (induction sur la forme des formules) que tout ensemble définit par une formule $\Pi_0 = \Sigma_0$ est décidable. Les sous-ensembles Δ_0 sont donc décidables.

Question 13. *Montrer que si une partie A de $\mathbb{N}^p$ est Σ_1 alors A est récursivement énumérable.*

On admettra que toute partie A de $\mathbb{N}^p$ décidable est Σ_1. On pourra aussi admettre qu'on peut [3] construire une bijection $\pi : \mathbb{N} \to \mathbb{N}^2$ telle que si l'on note $\pi(r) = (\pi_1(r), \pi_2(r))$, les relations $u = \pi_1(r)$ et $v = \pi_2(r)$ sont Δ_0.

Question 14. *Montrer la réciproque : une partie A de $\mathbb{N}^p$ est Σ_1 si elle est récursivement énumérable.*

Question 15. *En déduire que les ensembles Δ_1 de $\mathbb{N}^p$ sont exactement les ensembles décidables de $\mathbb{N}^p$.*

Fixons $p \geq 1$. On dit qu'un ensemble $U \subseteq \mathbb{N}^{p+1}$ est Σ_n-universel s'il est dans Σ_n et si pour tout $A \subseteq \mathbb{N}^p$ qui est dans Σ_n, il existe $i \in \mathbb{N}$ tel que $A = U_i$, où $U_i = \{(x_1, \ldots, x_p) | (i, x_1, \ldots, x_p) \in U\}$. On dit qu'un ensemble est Π_n universel, si on a la même propriété en remplaçant Σ_n par Π_n dans la phrase précédente.

On fixe un codage [4] des machines de Turing par les entiers, de telle sorte que l'on puisse parler de la machine de Turing numéro n.

Question 16. *Montrer que $U = \{(i, x_1, \ldots, x_p) |$ la machine de Turing i accepte $(x_1, \ldots, x_p)\}$ est Σ_1-universel.*

Question 17. *Montrer que pour tout $n \geq 1$, et pour tout $p \geq 1$, il existe un ensemble Σ_n-universel, et un ensemble Π_n-universel dans $\mathbb{N}^{p+1}$.*

3. C'est très facile.

4. Raisonnable, c'est-à-dire comme dans le cours, tel que l'on puisse bien retrouver chacun des ingrédients de la machine (programme, état initial, etc..) à partir de l'entier qui le code.

Question 18. *Montrer que les inclusions* $\Delta_n \subseteq \Sigma_n$ *et* $\Delta_n \subseteq \Pi_n$ *sont strictes pour* $n \geq 1$.

Question 19. *Montrer que chacun des niveaux* Σ_n *et* Π_n *pour* $n \geq 1$ *possède des problèmes complets. Donner un exemple de tels problèmes pour chacun des niveaux.*

Il est aussi possible de donner des caractérisations de chacun des niveaux de cette hiérarchie en utilisant des machines de Turing étendues (avec des oracles).