Fondements de l'informatique. Examen
Durée: 3h

Sujet proposé par Olivier Bournez
Version 4

Les 4 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.

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

Tous les graphes considérés sont non-orientés.

1 Machines de Turing

On fixe un alphabet fini Σ. Σ^* désigne l'ensemble des mots sur l'alphabet Σ.

Question 1. *Les problèmes suivants sont-ils décidables ? Justifier.*

- *Déterminer si le langage accepté par une machine de Turing M est co-fini : on dit qu'un langage est* co-fini *si son complémentaire (dans Σ^*) est fini.*
- *Déterminer si une machine de Turing accepte son entrée en utilisant au plus 2^{2^n} cases du ruban, où n est la longueur du mot en entrée.*
- *On fixe une fonction calculable $g : \mathbb{N} \to \mathbb{N}$. Déterminer si une machine de Turing accepte son entrée en utilisant au plus $2^{g(n)}$ cases du ruban, où n est la longueur du mot en entrée. Même question si l'on suppose que $g : \mathbb{N} \to \mathbb{N}$ est quelconque (c'est-à-dire, possiblement non-calculable).*

Une machine de Turing avec ruban semi-infini est une machine de Turing dont le programme est tel que la tête de lecture ne va jamais à gauche de sa position initiale.

Question 2. *Montrer que le problème de savoir si une machine de Turing semi-infinie accepte un mot w est indécidable.*

2 Espaces vectoriels

Soit E un $\mathbb{R}$-espace vectoriel. On dit que E est *ordonnable* s'il existe une relation d'ordre total $\leq$ sur E qui est compatible avec la structure d'espace vectoriel, c'est-à-dire :

- pour tout $x, x', y, y' \in E$, si $x \leq x'$ et $y \leq y'$ alors $x + y \leq x' + y'$;
- pour tout $x \in E$ et pour tout $\lambda \in \mathbb{R}$, si $x \geq 0$ et $\lambda \geq 0$ alors $\lambda x \geq 0$.

Question 3. *Construire un ensemble $\mathcal{F}_E$ de formules du calcul* ***propositionnel*** *sur une signature que vous préciserez tel que $\mathcal{F}_E$ est satisfiable si et seulement si E est ordonnable.*

Question 4. *Montrer qu'un $\mathbb{R}$-espace vectoriel E est ordonnable si et seulement si tous ses sous-espaces vectoriels de dimension finie*[1] *le sont.*

3 Chemins avec paires interdites

Etant donnés un graphe $G = (V, E)$, et un ensemble S de paires de sommets (u_1, v_1), (u_2, v_2), ..., $(u_k, v_k) \in V \times V$, on dit qu'un chemin de G évite les paires de S s'il visite au plus un sommet de chaque paire de S.

Question 5. *On considère le graphe G dont les sommets sont $V = \{s, t, u, v\}$ et les arêtes sont $E = \{(s, u), (s, v), (u, t), (v, t)\}$. Quels sont les chemins entre s et t qui évitent $S = \{(u, v)\}$?*

Question 6. *Construire un graphe à 5 sommets $V = \{s, t, u_1, u_2, u_3\}$ et un ensemble de paires S tels que :*

- *les chemins entre s et t qui évitent S passent nécessairement par l'un des trois sommets u_1, u_2, ou u_3 ;*
- *pour chaque sommet u_i pour $i \in \{1, 2, 3\}$, il y a au moins un chemin entre s et t qui évite S et passe par u_i.*

Question 7. *Démontrer que le problème suivant est NP-complet :*

- *Données : un graphe $G = (V, E)$, deux sommets $s, t \in V$, et un ensemble S de paires $(u_1, v_1), (u_2, v_2), \ldots, (u_k, v_k) \in V \times V$.*
- *Question : Existe-t-il un chemin entre s et t dans le graphe G qui évite les paires de S.*

On pourra utiliser la NP-complétude du problème 3-SAT.

4 Jouons aux dominos

Le problème de la correspondance de Post (PCP) est un problème de décision sur des *dominos* qui fut introduit par Emil Post en 1946. Il apparaît souvent dans des démonstrations d'indécidabilité : nous verrons un exemple plus bas concernant un problème sur les matrices.

Nous allons considérer plusieurs variantes de ce problème, et étudier leur difficulté.

On fixe un alphabet Σ.

On appelle "*domino*" un couple (U, V) où U et V sont des mots sur un alphabet Σ : on va représenter un tel couple graphiquement sous la forme $\boxed{\begin{array}{|c|}\hline U \\ \hline V \\ \hline \end{array}}$.

Etant donné une suite finie S de dominos, le problème est de savoir si l'on peut poser ces dominos l'un à la suite de l'autre (dans n'importe quel ordre) de telle sorte que le mot qui apparaît en haut soit le même que le mot qui apparaît en bas : on appelle cela une *correspondance*.

On souhaite savoir s'il existe un algorithme qui, étant donnée une suite finie S de dominos, décide si il existe une correspondance pour cette suite : on appelle cela le problème de correspondance de Post.

4.1 Sans répétition

Supposons que l'on a le droit d'utiliser chaque domino de S au plus une fois[2].

Par exemple, si l'on part de l'ensemble

$$S = \left\{ \boxed{\begin{array}{|c|}\hline ba \\ \hline ac \\ \hline \end{array}}, \boxed{\begin{array}{|c|}\hline a \\ \hline ab \\ \hline \end{array}}, \boxed{\begin{array}{|c|}\hline ca \\ \hline a \\ \hline \end{array}} \right\}$$

1. On rappelle qu'un sous-espace vectoriel de dimension finie de E est le plus petit sous-espace vectoriel engendré par (i.e. qui contient) un nombre fini d'éléments de E.
2. On n'interdit pas cependant qu'un même domino puisse avoir plusieurs occurrences dans S.

on peut construire la correspondance suivante :

$$\left[\begin{array}{|c|}\hline a \\ \hline ab \\ \hline\end{array}\right]\left[\begin{array}{|c|}\hline ba \\ \hline ac \\ \hline\end{array}\right]\left[\begin{array}{|c|}\hline ca \\ \hline a \\ \hline\end{array}\right] :$$

on a bien en haut et en bas le même mot, à savoir *abaca*.

Si cela aide de le dire très formellement : étant donné une suite de dominos [3]

$$S = \left\{\left[\begin{array}{|c|}\hline U_1 \\ \hline V_1 \\ \hline\end{array}\right], \left[\begin{array}{|c|}\hline U_2 \\ \hline V_2 \\ \hline\end{array}\right], \ldots, \left[\begin{array}{|c|}\hline U_p \\ \hline V_p \\ \hline\end{array}\right]\right\}, \tag{1}$$

le problème de la correspondance de Post **sans répétition** consiste à déterminer s'il existe une suite non vide d'indices $i_1, i_2, \ldots, i_k$ dans $\{1, 2, \ldots, p\}$ telle que

$$U_{i_1} U_{i_2} \ldots U_{i_k} = V_{i_1} V_{i_2} \ldots V_{i_k}.$$

$$\left(\text{ graphiquement : } \left[\begin{array}{|c|}\hline U_{i_1} \\ \hline V_{i_1} \\ \hline\end{array}\right]\left[\begin{array}{|c|}\hline U_{i_2} \\ \hline V_{i_2} \\ \hline\end{array}\right]\left[\begin{array}{|c|}\hline U_{i_3} \\ \hline V_{i_3} \\ \hline\end{array}\right] \ldots \left[\begin{array}{|c|}\hline U_{i_k} \\ \hline V_{i_k} \\ \hline\end{array}\right] \right)$$

avec $i_k \neq i_l$ pour $k \neq l$.

Question 8. *Montrer que le problème de la correspondance de Post sans répétition est décidable.*

On établira dans la section 4.6 qu'il est NP-complet.

4.2 Avec répétitions

Jusqu'à la section 4.6, on s'autorise à utiliser plusieurs fois un même domino de la suite S.

Par exemple, si l'on part de la suite

$$S_1 = \left\{\left[\begin{array}{|c|}\hline b \\ \hline ca \\ \hline\end{array}\right], \left[\begin{array}{|c|}\hline a \\ \hline ab \\ \hline\end{array}\right], \left[\begin{array}{|c|}\hline ca \\ \hline a \\ \hline\end{array}\right], \left[\begin{array}{|c|}\hline abc \\ \hline c \\ \hline\end{array}\right]\right\}$$

on peut construire la correspondance suivante :

$$\left[\begin{array}{|c|}\hline a \\ \hline ab \\ \hline\end{array}\right]\left[\begin{array}{|c|}\hline b \\ \hline ca \\ \hline\end{array}\right]\left[\begin{array}{|c|}\hline ca \\ \hline a \\ \hline\end{array}\right]\left[\begin{array}{|c|}\hline a \\ \hline ab \\ \hline\end{array}\right]\left[\begin{array}{|c|}\hline abc \\ \hline c \\ \hline\end{array}\right] :$$

on a bien en haut et en bas le même mot, à savoir *abcaaabc* (et on a utilisé plusieurs fois un même domino).

Tous les ensembles S ne possèdent pas de correspondance :

Question 9. *On part cette fois de la suite*

$$S_2 = \left\{\left[\begin{array}{|c|}\hline abc \\ \hline ab \\ \hline\end{array}\right], \left[\begin{array}{|c|}\hline ca \\ \hline a \\ \hline\end{array}\right], \left[\begin{array}{|c|}\hline acc \\ \hline ba \\ \hline\end{array}\right]\right\}$$

Est-il possible d'obtenir une correspondance ?

Si cela aide de le dire formellement : étant donné un ensemble de dominos

$$S = \left\{\left[\begin{array}{|c|}\hline U_1 \\ \hline V_1 \\ \hline\end{array}\right], \left[\begin{array}{|c|}\hline U_2 \\ \hline V_2 \\ \hline\end{array}\right], \ldots, \left[\begin{array}{|c|}\hline U_p \\ \hline V_p \\ \hline\end{array}\right]\right\},$$

le problème de la correspondance de Post (avec répétitions) consiste à déterminer s'il existe une suite non vide d'indices $i_1, i_2, \ldots, i_k$ dans $\{1, 2, \ldots, p\}$ telle que

$$U_{i_1} U_{i_2} \ldots U_{i_k} = V_{i_1} V_{i_2} \ldots V_{i_k}$$

$$\left(\text{ graphiquement : } \left[\begin{array}{|c|}\hline U_{i_1} \\ \hline V_{i_1} \\ \hline\end{array}\right]\left[\begin{array}{|c|}\hline U_{i_2} \\ \hline V_{i_2} \\ \hline\end{array}\right]\left[\begin{array}{|c|}\hline U_{i_3} \\ \hline V_{i_3} \\ \hline\end{array}\right] \ldots \left[\begin{array}{|c|}\hline U_{i_k} \\ \hline V_{i_k} \\ \hline\end{array}\right] \right).$$

3. Pour les plus puristes, et en lien avec la note de bas de page 2. (page précédente) : la notation de la suite S dans l'écriture sous forme ensembliste (1) est possiblement la notation d'un multi-ensemble plutôt qu'un ensemble : par rapport à un ensemble, on autorise un même élement à apparaitre deux fois ou plus dans un multi-ensemble.

4.3 Réduction à la variante modifiée

On s'intéresse à la variante suivante, que l'on appelle problème de correspondance de Post **modifié** : on impose que $i_1 = 1$, c'est-à-dire, on impose le premier domino.

Si cela aide de le dire très formellement : étant donné un ensemble de dominos

$$S = \left\{ \boxed{\begin{array}{c} U_1 \\ \hline V_1 \end{array}}, \boxed{\begin{array}{c} U_2 \\ \hline V_2 \end{array}}, \ldots, \boxed{\begin{array}{c} U_p \\ \hline V_p \end{array}} \right\},$$

le problème de la correspondance de Post **modifié** consiste à déterminer s'il existe une suite non vide d'indices $i_2, i_3, \ldots, i_k$ dans $\{1, 2, \ldots, p\}$ telle que

$$U_1 U_{i_2} \ldots U_{i_k} = V_1 V_{i_2} \ldots V_{i_k}.$$

$$\left(\text{ graphiquement : } \boxed{\begin{array}{c} U_1 \\ \hline V_1 \end{array}} \boxed{\begin{array}{c} U_{i_2} \\ \hline V_{i_2} \end{array}} \boxed{\begin{array}{c} U_{i_3} \\ \hline V_{i_3} \end{array}} \cdots \boxed{\begin{array}{c} U_{i_k} \\ \hline V_{i_k} \end{array}} \right).$$

Soit $u = u_1 u_2 \ldots u_n$ un mot de longueur n sur l'alphabet Σ. Soient $*$ et $\diamond$ deux nouvelles lettres : $* \notin \Sigma$, $\diamond \notin \Sigma$.

On définit $*u$, $u*$ et $*u*$ comme les mots suivants :

$$*u = *u_1 * u_2 * u_3 \cdots * u_n$$

$$u* = u_1 * u_2 * u_3 \cdots * u_n*$$

$$*u* = *u_1 * u_2 * u_3 \cdots * u_n*$$

En d'autres termes, $*u$ (respectivement : $u*$, $*u*$) ajoute le symbole $*$ avant (resp. après, avant et après) chaque lettre du mot u.

Question 10. *On considère*

$$S_1^* = \left\{ \boxed{\begin{array}{c} *a \\ \hline *a*b* \end{array}}, \boxed{\begin{array}{c} *b \\ \hline c*a* \end{array}}, \boxed{\begin{array}{c} *a \\ \hline a*b* \end{array}}, \boxed{\begin{array}{c} *c*a \\ \hline a* \end{array}}, \boxed{\begin{array}{c} *a*b*c \\ \hline c* \end{array}}, \boxed{\begin{array}{c} *\diamond \\ \hline \diamond \end{array}} \right\}$$

(Observer comment S_1^* s'obtient à partir de S_1 en début de la section 4.2 en appliquant les opérations ci-dessus en haut et en bas sur les dominos).

Expliquer comment les correspondances[4] *de S_1^* s'obtiennent à partir de celles de S_1 et réciproquement.*

Question 11. *Démontrer que le problème de Post modifié se réduit au problème de Post.*

4.4 Indécidabilité de la variante modifiée

Question 12. *Démontrer que le problème de la correspondance de POST modifié est indécidable.*

Etant donnés une machine de Turing M semi-infinie et un mot w on cherchera à construire un ensemble S_M de dominos tel qu'une correspondance de S_M décrit un calcul de M sur l'entrée w. On pourra inclure dans S_M les dominos

$$\boxed{\begin{array}{c} \# \\ \hline \#q_0 w \# \end{array}}, \boxed{\begin{array}{c} a \\ \hline a \end{array}}, \boxed{\begin{array}{c} \# \\ \hline B\# \end{array}}, \boxed{\begin{array}{c} aq_a \\ \hline q_a \end{array}}, \boxed{\begin{array}{c} q_a a \\ \hline q_a \end{array}}, \boxed{\begin{array}{c} q_a \#\# \\ \hline \# \end{array}}$$

et d'autres dominos bien choisis que l'on décrira : w est un mot, $\#$,a une lettre, B le symbole de blanc, q_0 et q_a l'état initial et acceptant de M.

Par conséquent, le problème de la correspondance de Post est indécidable.

4. On parle bien des correspondances "normales", i.e. au sens des questions précédentes, pas nécessairement "modifées" comme dans le texte qui précède.

4.5 Application : Problème de matrices

Un ensemble H de matrices 3×3 à coefficients entiers est dit *mortel* si la matrice nulle peut s'obtenir comme un produit d'un nombre fini de matrices de cet ensemble (on s'autorise à utiliser plusieurs fois la même matrice de l'ensemble H dans le produit).

Le but de cette partie est de prouver qu'il est indécidable de déterminer si un ensemble H de matrices est mortel.

On considere un ensemble H de matrices 3×3 qui contient les matrices

$$S = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad T = \begin{bmatrix} 1 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

ainsi que des matrices de la forme

$$W_j = \begin{bmatrix} p_j & 0 & 0 \\ 0 & r_j & 0 \\ q_j & s_j & 1 \end{bmatrix} \tag{2}$$

pour certains entiers p_j, q_j, r_j, s_j.

On note $[x, y, z]$ pour un vecteur ligne.

En observant que

$$[a, b, c]S = a[1, 0, 1]$$

$$[a, b, c]T = (a - b)[1, -1, 0]$$

et que toutes les matrices W_j sont inversibles, un raisonnement assez simple (sur les types de produits possibles avec un tel ensemble H et sur les images et noyaux de S et T) démontre le fait suivant que l'on admettra : *une condition nécessaire et suffisante pour obtenir un produit nul avec les matrices de l'ensemble H est qu'il existe un produit X des matrices W_j tel que*

$$[1, 0, 1]X = [h, h, 1]$$

pour un $h > 0$ (et dans ce cas $SXT = 0$).

Illustrons une relation entre la multiplication de matrices et la concaténation de mots par un exemple :

$$[123, 12, 1] \begin{bmatrix} 1000 & 0 & 0 \\ 0 & 100 & 0 \\ 223 & 32 & 1 \end{bmatrix} = [123223, 1232, 1]$$

Question 13. *Etant donnée une paire de mots $< U, V >$, définir une matrice $W(U, V)$ de la forme W_j (i.e. de la forme (2)) telle si X, Y, U, V sont des mots sur l'alphabet $\{1, 2, 3\}$, alors*

$$[X, Y, 1]W(U, V) = [X.U, Y.V, 1]$$

où . désigne la concaténation, et les mots sont interprétés comme des entiers de la façon évidente.

Question 14. *Démontrer qu'il est indécidable de savoir si un ensemble H de matrices est mortel.*

4.6 Retour sur les versions sans répétition

On revient sur la variante du problème de la correspondance de Post **sans répétition**, considérée dans la section 4.1 (rappel : on peut bien avoir plusieurs fois un même domino dans la suite S, mais chaque domino ne peut être utilisé qu'une fois dans une correspondance).

Question 15. *Montrer que le problème de correspondance de Post sans répétition est NP-complet.*

Question 16. *On ne s'autorise pas les répétitions de matrices. Montrer qu'il est décidable de savoir si un ensemble H de matrices est mortel sans répétition. Montrer que le problème est NP-complet.*