Fondements de l'informatique: Examen
Durée: 3h

Sujet proposé par Olivier Bournez
Version 5
(corrigé)

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.

Il est possible d'avoir la note maximale sans avoir répondu à toutes les questions. Les questions avec un barème plus élevé sont indiquées par le symbole $(*)$.

On pourra utiliser les résultats et 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 JAVA, C ou OCAML.

1 Calculabilité

Question 1. *Soient A et B deux langages récursivement énumérables sur l'alphabet Σ tels que*[1] *$A \cup B = \Sigma^*$ et $A \cap B$ est un ensemble récursif. Démontrer que A et B sont récursifs.*

Solution : Soit $M_{A\cap B}$ la machine de Turing qui décide $A \cap B$, et M_A la machine qui reconnait A, et M_B celle qui reconnaît B. Le langage A est récursif car il est décidé par la machine qui sur un mot w, simule $M_{A\cap B}$ sur w et accepte si $M_{A\cap B}$ accepte, et sinon fait pour $t = 1, 2, \ldots etc$,

— simule t étapes de M_A et s'arrête en acceptant si M_A accepte pendant ces t étapes ;

— simule t étapes de M_B et s'arrête en refusant si M_B accepte pendant ces t étapes

La machine s'arrête nécessairement, car soit $w \in A \cap B$, soit $w \notin A \cap B$, et alors $w \in A$ ou $w \in B$, et par conséquent pour un certain t, soit M_A accepte, soit M_B accepte. Par ailleurs, la réponse est toujours la bonne par construction. □

Question 2. *Parmi les problèmes suivants, lesquels sont décidables ? lesquels sont indécidables*[2] *?*

(Dans cette question et la suivante, les machines de Turing travaillent sur l'alphabet $\Sigma = \{a, b, \ldots, z\}$.)

1. Σ^* désigne l'ensemble des mots sur l'alphabet Σ.
2. Rappel : indécidable signifie non-décidable.

1. *Déterminer si un programme JAVA contient la chaîne de caractères "examen".*
2. *Déterminer si un programme JAVA affiche la chaîne de caractères "examen" lorsqu'il est exécuté.*
3. *Déterminer si le langage accepté par une machine de Turing contient le mot "examen".*
4. *Déterminer si une machine de Turing s'arrête sur toute entrée.*

Solution :

Le problème de la question 1. est décidable. Il suffit de rechercher la chaîne "examen" dans le texte du programme.

Le problème de la question 2. est indécidable : le problème de l'arrêt des machines de Turing se réduit à ce problème. En effet, étant données une machine M et une entrée w, on peut construire un programme JAVA qui simule M sur w et si M accepte w affiche "examen". Ce programme JAVA affiche examen si et seulement si M accepte w.

Le problème de la question 3. est indécidable. C'est une application directe du théorème de Rice.

Le problème de la question 4. est indécidable : le problème de l'arrêt des machines de Turing se réduit à ce problème : en effet, étant donnée une machine M et une entrée w, on peut construire une machine M' qui sur une entrée u, simule M sur w et si M accepte u. M' accepte toutes les entrées si et seulement si M accepte w. □

Question 3. *Parmi les problèmes 1., 2., 3. de la question précédente, lesquels sont récursivement énumérables ou non-récursivement énumérables ? Justifier votre réponse.*

(∗) *Même question pour 4. (On pourra utiliser pour 4. le complémentaire du problème de l'arrêt des machines de Turing).*

Solution : Tout problème décidable est récursivement énumérable. Donc 1. est récursivement énumérable. Le problème 2. est récursivement énumérable, car il suffit de simuler le programme JAVA, et de s'arrêter dès qu'il affiche "examen" en acceptant.

Le problème 3. est récursivement énumérable, car il suffit de simuler la machine de Turing sur le mot "examen" et accepter si la machine simulée accepte.

Le problème 4. n'est pas récursivement énumérable, car s'il était récursivement énumérable, alors le complémentaire du problème de l'arrêt des machines de Turing le serait.

En effet, le complémentaire du problème de l'arrêt des machines de Turing se réduit à ce problème : étant donnée une machine de Turing M, et une entrée w, on peut construire la machine de Turing M_w qui sur une entrée t accepte si M n'accepte pas w en un temps inférieur à (l'entier codé par) t, et boucle sinon. On a en effet que M n'accepte pas w ssi $L(M_w)$ contient tous les entiers ssi M_w est récursif. □

2 Un peu de définissabilité sur les graphes

On considère une signature Σ qui contient un prédicat binaire R, et un symbole binaire d'égalité $=$.

Question 4. *Rappeler pourquoi les modèles (synonyme : structures) sur cette signature correspondent à des graphes orientés*[3]*.*

Solution : Un modèle correspond à un ensemble sous-jacent, que l'on peut voir comme des sommets, et à une interprétation de la relation R, que l'on peut voir comme les arrêtes. □

3. Comme dans le cours, ce que nous appelons graphes ne sont pas des multigraphes : il y a au plus une arête orientée de s vers t pour chaque couple de sommet s, t.

On rappelle qu'un sommet s d'un graphe est de degré sortant k, s'il possède exactement k voisins distincts e_1, e_2, ..., e_k avec une arête orientée[4] de s vers e_i pour $1 \leq i \leq k$.

Question 5. *Ecrire une formule F_3 sur la signature Σ qui caractérise les graphes de degré sortant supérieur ou égal à 3 : la formule est vraie dans un graphe G si et seulement si G a tous ses sommets de degré sortant supérieur ou égal à 3.*

Solution : On écrit $\forall x \exists y_1 \exists y_2 \exists y_3\ R(x, y_1) \wedge R(x, y_2) \wedge R(x, y_3) \wedge y_1 \neq y_2 \wedge y_1 \neq y_3 \wedge y_2 \neq y_3$.
($y_i \neq y_j$ est bien entendu $\neg y_i = y_j$). □

Un graphe G est de degré sortant fini s'il existe un entier k tel que tous les sommets soient de degré sortant inférieur ou égal à k.

Question 6. *Montrer qu'il n'est pas possible de caractériser les graphes de degré sortant fini : il n'y a pas de formule F, telle que F soit vraie dans un graphe G si et seulement si G est de degré sortant fini.*

Solution : Cela découle du théorème de compacité.

En effet, par l'absurde. Supposons qu'il existe une telle formule Φ. On considère $\mathcal{T}$ comme l'union de la formule Φ et de toutes les formules F_k, où F_k est la formule $\forall x \exists y_1 \exists y_2 \ldots \exists y_k\ R(x, y_1) \wedge R(x, y_2) \wedge \cdots \wedge R(x, y_k) \wedge y_1 \neq y_2 \wedge y_1 \neq y_3 \wedge \ldots y_1 \neq y_k \wedge y_2 \neq y_3 \wedge y_2 \neq y_4 \wedge y_2 \neq y_k \wedge \cdots \wedge y_{k-1} \neq y_k$.

Toute partie de $\mathcal{T}$ est consistante : choisir un graphe de degré sortant suffisant impliquant les sommets mentionnés dans le sous-ensemble de $\mathcal{T}$. Par le théorème de compacité, $\mathcal{T}$ est donc consistant, ce qui signifie qu'il possède un modèle (un graphe) dont le degré sortant est plus grand que k pour tout k (c'est ce qu'exprime les formules F_k). Or $\mathcal{T}$ contient Φ et donc contredit Φ. Absurde. □

3 NP-complétude

L'objectif de cet exercice est de prouver que le problème de décision $Max2SAT$ suivant est NP-complet :

- **Donnée**: Un ensemble F' de clauses d'au plus 2 littéraux, un entier k ;
- **Réponse**: Décider s'il existe une assignation $x_1, \ldots, x_n \in \{0, 1\}$ des variables telle que au moins k clauses de F' s'évaluent à vrai pour cette valeur des variables $x_1, \ldots, x_n$.

Question 7. *Soit la clause $C = x \vee y \vee z$. Nous construisons l'ensemble de clauses ϕ_C défini comme l'ensemble suivant (qui contient 10 clauses) :*

$$\phi_C = \{(x), (y), (z), (p_C), (\neg x \vee \neg y), (\neg y \vee \neg z), (\neg z \vee \neg x), (x \vee \neg p_C), (y \vee \neg p_C), (z \vee \neg p_C)\}$$

(où l'on a introduit un nouveau[5] littéral p_C). Déterminer le nombre maximum de clauses simultanément satisfiables dans l'ensemble ϕ_C dans le cas où la clause C est satisfaite. (On pourra raisonner selon le nombre de littéraux égaux à vrai dans la clause C.)

Même question lorsque la clause C n'est pas satisfaite.

Solution : Considérons le cas où la clause C est satisfaite. Nous allons énumérer les différents cas.

1. Supposons que tous les littéraux de C sont égaux à vrai. Les 3 premières clauses et les 3 dernières de ϕ_C sont satisfaites. De plus, les cinquième, sixième, septième clauses de ϕ_C ne sont pas satisfaites. Ensuite, la clause (p_C) est satisfaite si $p_C = 1$. Le nombre maximum de clauses satisfaites est égale à 7.

4. Une arête orientée s'appelle aussi un arc.
5. Distinct de x, y, z.

2. Supposons qu'un seul littéral de C est égal à faux. Sans perte de généralité supposons que cela soit $z = 0$. Alors les clauses (x), (y), $(\neg y \vee \neg z)$, $(\neg z \vee \neg x)$, $(x \vee \neg p_C)$, $(y \vee \neg p_C)$ sont satisfaites. Quelle que soit la valeur de p_C, seule une de ces deux clauses (p_C) et $(z \vee \neg p_C)$ peut être satisfaite. Le nombre maximum de clauses satisfaites est égale à 7.
3. Supposons que deux littéraux de C sont égaux à faux. Sans perte de généralité supposons que cela soit $z = 0$ et $y = 0$. Alors les clauses (x), $(\neg x \vee \neg y)$, $(\neg y \vee \neg z)$, $(\neg z \vee \neg x)$, $(x \vee \neg p_C)$ sont satisfaites. En prennant $p_C = 0$, le nombre maximun de clauses satisfaites parmi les 3 clauses (p_C) et $(y \vee \neg p_C)$ et $(z \vee \neg p_C)$ vaut 2. Le nombre maximum de clauses satisfaites simultanément satisfiables est égale à 7.

Donc si la clause C est satisfaite, alors le nombre maximum de clauses satisfaites de ϕ_C est égal à 7.

Si la clause C n'est pas satisfaite : les trois littéraux doivent être à faux. Cela signifie que les 3 premières clauses de ϕ_C ne sont pas satisfaites. Les cinquième, sixième, septième clauses de ϕ_C sont satisfaites. Ensuite, la valeur des autres clauses dépendent du littéral p_C. Afin de maximiser le nombre de clauses, il faut que $p_C = 1$.

Donc si la clause C n'est pas satisfaite, alors le nombre maximum de clauses satisfaites de ϕ_C est égale à 6.

□

Question 8. *Montrer que le problème Max2SAT est dans NP.*

Solution : Le problème Max2SAT est clairement dans NP, car la donnée de la valeur des variables constitue un certificat vérifiable en temps polynomial : il suffit de vérifier qu'au moins k clauses de F' s'évaluent à vrai pour cette valeur des variables. □

Question 9. *Montrer que le problème Max2SAT est NP-complet.*

Solution : Il reste à montrer qu'il est plus difficile qu'un problème NP-complet.

Nous allons réduire 3SAT à Max2SAT de la façon suivante. Etant donnée une instance I de 3SAT, on construit une instance I' de Max2SAT de la façon suivante : pour chaque clause $C_i = (x_1 \vee x_2 \vee x_3)$ de F, on rajoute les 10 clauses précédentes à F' où x_1, x_2 et x_3 remplacent x, y et z. La variable p_C est remplacée par une variable supplémentaire p_{C_i} associée à la clause C_i. Ainsi, I' a $10 \cdot m$ clauses si F en possède m et $k = 7m$. La réduction proposée est bien polynomiale.

D'après les questions précédentes, au moins k clauses de F' est satisfiable ssi F est satisfiable.

□

4 Modèles de Herbrand

Soit $\Sigma = (\mathcal{C}, \mathcal{F}, \mathcal{R})$ une signature d'un langage du premier ordre.

L'*univers de Herbrand* de $\mathcal{L} = \mathcal{C} \cup \mathcal{F}$ est l'ensemble U_H des termes clos construits à partir des symboles de constantes de $\mathcal{C}$ et des symboles de fonctions de $\mathcal{F}$.

Une structure sur la signature Σ est une *structure de Herbrand* si son domaine [6] est l'univers de Herbrand, et si chaque terme clos est interprété par lui même.

Pour une signature Σ donnée, il n'y a qu'un seul univers de Herbrand, mais sur cet univers de Herbrand, on peut définir possiblement de nombreuses structures de Herbrand H, en faisant varier l'interprétation des symboles de relation.

6. Rappel : "domaine" est un synonyme pour "ensemble de base".

Pour déterminer une interprétation, il faut et il suffit de déterminer pour chaque formule atomique close sur la signature Σ si elle est vraie ou fausse :

Appelons *base de Herbrand* de Σ l'ensemble des formules atomiques closes sur la signature Σ : c'est-à-dire, les formules de la forme $R(t_1, \ldots, t_n)$ avec $R \in \mathcal{R}$, $t_1, \ldots, t_n \in U_H$.

Une structure de Herbrand est en effet parfaitement définie par un sous ensemble I_H de la base de Herbrand :

— à chaque structure de Herbrand H sur Σ correspond le sous-ensemble I_H de la base de Herbrand où I_H est défini comme les formules atomiques qui sont vraies dans H.

— Réciproquement, chaque sous-ensemble I_H de la base de Herbrand définit bien une structure de Herbrand, en considérant la structure de Herbrand où l'on suppose vraies les formules atomiques de I_H et fausses les autres.

Question 10. *On considère le cas où $\mathcal{C} = \{a\}$ (il n'y a qu'un seul symbole de constante a) et $\mathcal{F} = \emptyset$ (il n'y a pas de symbole de fonction) et $\mathcal{R} = \{P, Q\}$ (il y a deux symboles de relations P et Q d'arité 1).*

Décrire les 4 structures de Herbrand de la signature $\Sigma = (\mathcal{C}, \mathcal{F}, \mathcal{R})$ en donnant les ensembles I correspondants.

Solution : L'univers de Herbrand contient que $\{a\}$. Le seul terme clos est donc la constante a. Il y a exactement 4 structures de Herbrand : pour la première, $P(a)$ et $Q(a)$ sont faux : $I_1 = \emptyset$. Pour la seconde, $P(a)$ est vraie, et $Q(a)$ est fausse : $I_2 = \{P(a)\}$. Pour la troisième, $P(a)$ est fausse, et $Q(a)$ est vraie : $I_3 = \{Q(a)\}$. Dans la quatrième, $P(a)$ est vraie, et $Q(a)$ est vraie. $I_4 = \{P(a), Q(a)\}$. □

On fixe une signature Σ, où l'ensemble $\mathcal{C}$ des constantes est non-vide (afin que l'univers de Herbrand soit non-vide).

On dit qu'un ensemble $\mathcal{T}$ de formules universelles [7] (sur une signature Σ) *possède un plus petit modèle de Herbrand*, s'il y a une structure de Herbrand H (sur la signature Σ) qui est un modèle de $\mathcal{T}$, et si pour toute structure de Herbrand L (sur la signature Σ) qui est modèle de $\mathcal{T}$, on a I_H inclus dans I_L.

Tout ensemble $\mathcal{T}$ de formules universelles ne possède pas nécessairement un plus petit modèle de Herbrand :

Question 11. *En effet, montrer que pour $\mathcal{T}$ réduite à la formule $\forall x\ (P(x) \vee Q(x))$ pour la signature de la question précédente, il n'y a pas de plus petit modèle de Herbrand.*

Solution : Les modèles de $\mathcal{T}$ sont $I_2 = \{P(a)\}$, $I_3 = \{Q(a)\}$, $I_4 = \{P(a), Q(a)\}$ (car I_1 ne correspond pas à un modèle de $\forall x\ (P(x) \vee Q(x))$). Il n'y a pas de sous-ensemble de ces trois ensembles qui soit inclus dans les trois et qui corresponde à une structure de Herbrand (l'ensemble vide est le seul candidat et correspond à I_1 que nous venons de dire ne pas satisfaire $\forall x\ (P(x) \vee Q(x))$). □

Un *littéral* est une formule atomique ou la négation d'une formule atomique : c'est-à-dire de la forme $R(t_1, \ldots, t_r)$ (littéral positif) ou de la forme $\neg R(t_1, \ldots, t_r)$ (littéral négatif), où $R \in \mathcal{R}$ est un symbole de relation, et r est son arité.

Question 12. *Une* clause programme PROLOG [8] *est une formule universelle de la forme*

$$\forall x_1 \forall x_2 \ldots \forall x_p (L_1 \vee \cdots \vee L_n)$$

où les L_i sont des littéraux, et exactement un de ces littéraux est positif.

7. C'est-à-dire de la forme $\forall x_1 \forall x_2 \ldots \forall x_n \psi$, où ψ est sans quantificateur.

8. PROLOG est un langage de programmation logique.

(∗) *Montrer que tout ensemble $\mathcal{T}$ de clauses programmes PROLOG possède un plus petit modèle de Herbrand.*
Indication : on pourra considérer

$$I_M = \bigcap_{H \text{ structure de Herbrand modèle de } \mathcal{T}} I_H.$$

Solution : On prouve que l'intersection I_M de toutes les I_H pour un H correspondant à une structures de Herbrand est un ensemble I_M qui définit une structure de Herbrand M minimale.

Il suffit de prouver que pour chaque clause PRLOG C_i de $\mathcal{T}$, C_i est valide dans M.

Soit C_i est de la forme $\forall x_1 \dots \forall x_p\ A$ où A est positif : alors toutes les instances closes de A sont vraies dans tous les modèles de Herbrand, et donc elles sont dans toutes les I_H et donc aussi dans I_M, et donc sont vraies en M.

Sinon C_i est de la forme $\forall x_1 \dots \forall x_p(\neg A_1 \vee \neg A_2 \vee \dots \vee \neg A_n \vee A)$. On pose $C_i' = (\neg A_1 \vee \neg A_2 \vee \dots \vee \neg A_n \vee A)$.

Pour v une fonction de $\{x_1, \dots, x_p\}$ dans U_H, et pour $B \in \{A, A_1, \dots, A_n\}$, on note $v^*(B)$ l'atome clos en subtituant $v(x_i)$ à x_i dans B.

— soit il y a un A_j avec $v^*(A_j) \notin I_M$: dans ce cas, $\neg A_j$ s'évalue en vrai pour cette valuation, et donc C_i' s'évalue en vrai.

— soit pour tout A_j on a $v^*(A_j) \in I_M$, et donc pour tout I_H, on a $v^*(A_j) \in I_H$, et comme H est une structure de Herbrand, on doit avoir $v^*(A) \in I_H$, et donc A s'évalue en vrai, et donc C_i' s'évalue en vrai.

□

5 Une excursion en analyse non-standard

On admettra l'existence [9] d'une fonction [10] [11] [12] μ définie sur l'ensemble des parties de $\mathbb{N}$, à valeurs dans $\{0, 1\}$, telle que :

0. $\mu(X) = 0$ ou $\mu(X) = 1$ pour toute partie X de $\mathbb{N}$;
1. $\mu(\mathbb{N}) = 1$;
2. $\mu(X) = 0$ pour toute partie finie X de $\mathbb{N}$;
3. $\mu(X \cup Y) = \mu(X) + \mu(Y)$ pour toutes parties X, Y de $\mathbb{N}$ telles que $X \cap Y = \emptyset$.

On peut déduire de ces propriétés que :

4. $\mu(X) = 0$ si et seulement si $\mu(\mathbb{N} - X) = 1$.
5. $\mu(X) = 1$ et $\mu(Y) = 1$ implique $\mu(X \cap Y) = 1$.

5.1 L'ensemble des hyperréels

On construit l'ensemble ${}^*\mathbb{R}$ des hyperréels à partir de l'ensemble des réels $\mathbb{R}$: on considère une relation d'équivalence $\sim$ bien particulière sur l'ensemble des suites de réels.

Concrètement : si (x_i) et (y_i) sont des suites de réels, on note $(x_i) \sim (y_i)$ lorsque $\mu(\{i | x_i = y_i\}) = 1$.

9. On a besoin de l'axiome du choix pour construire cette fonction.
10. μ est appelée une mesure, ou aussi ultrafiltre non-principal.
11. Ce n'est pas une mesure au sens de la théorie de la mesure, car on ne la suppose pas σ-additive si vous connaissez la théorie de la mesure.
12. Si cela peut aider, sans que cela soit certain de vous aider si vous ne comprenez pas cette intuition qui n'est pas indispensable : l'intuition de toute la suite est que $\mu(\mathbb{N}) = 1$ signifie que X est une "grosse" partie de $\mathbb{N}$, et qu'alors une propriété vraie sur X est vraie "presque partout" sur les entiers (i.e. pour "presque" tous les entiers). Les propriétés 0., 1.,2.,3., qui garantissent aussi 4. et 5., visent alors à faire fonctionner cette intuition dans ce qui suit.

Question 13. *Démontrer que $\sim$ est une relation d'équivalence : autrement dit, démontrer que pour toutes suites de réels (x_i), (y_i) et (z_i), on a*

1. $(x_i) \sim (x_i)$.
2. $(x_i) \sim (y_i)$ *implique* $(y_i) \sim (x_i)$.
3. $(x_i) \sim (y_i)$ *et* $(y_i) \sim (z_i)$ *implique* $(x_i) \sim (z_i)$.

Solution : Le point 1. découle du fait que $\mu(\mathbb{N}) = 1$. Le point 2. du fait que $=$ est symétrique. Le point 3. du fait que si $\mu(\{i|x_i = y_i\}) = 1$ et $\mu(\{i|y_i = z_i\}) = 1$ alors $\mu(\{i|x_i = y_i = z_i\}) = 1$ par la propriété 5. □

Lorsque (x_i) est une suite, $< x_i >$ désigne sa classe d'équivalence : si vous préférez, $< x_i >$ est l'ensemble des suites (y_i) avec $(x_i) \sim (y_i)$.

On définit alors ${}^*\mathbb{R}$, l'ensemble des hyperréels, comme l'ensemble [13] des classes d'équivalence :

$$ {}^*\mathbb{R} = \{< x_i > |(x_i) \text{ suite de réels}\}. $$

On peut considérer que les réels sont des hyperréels particuliers : $a \in \mathbb{R}$ correspond à ${}^*a \in {}^*\mathbb{R}$, où ${}^*a =< a >$ (i.e. est la classe d'équivalence de la suite constante dont tous les termes sont égaux à a).

On définit alors les opérations ${}^*+$ (addition), ${}^*\times$ (multiplication), et la relation ${}^* \leq$ à partir de celles sur $\mathbb{R}$ selon le principe suivant :

— $< x_i > {}^*+ < y_i >=< x_i + y_i >$;
— $< x_i > {}^*\times < y_i >=< x_i \times y_i >$;
— $< x_i > \ {}^*\leq < y_i >$ si et seulement si $\mu(\{i|x_i \leq y_i\}) = 1$;

On définit la valeur absolue par ${}^*| < x_i > | =< |x_i| >$.

(On peut se convaincre que tout cela est bien défini [14] et étend bien les opérations et l'ordre sur $\mathbb{R}$ à ${}^*\mathbb{R}$.).

5.2 Hyperréels infiniments petits

Un premier résultat fondamental est l'existence d'infiniments petits non-nuls : un hyperréel $x \in {}^*\mathbb{R}$ est dit *infiniment petit non nul* si $\neg x = {}^*0$ et ${}^*|x| \ {}^*\leq {}^*r$ pour tout réel $r \in \mathbb{R}$.

Question 14. *Démontrer que si (x_i) est une suite de réels non-nuls de limite 0, alors l'hyperréel $< x_i >$ est un infiniment petit non nul.*

Solution :

(On suppose bien entendu que $r > 0$ dans l'énoncé au dessus, sinon l'énoncé n'a pas de sens.)

On pose $a =< x_i >$. On a $a \neq {}^*0$ car $\mu(\{i|x_i = 0\}) = \mu(\emptyset) = 0$. Soit r un réel > 0. L'ensemble $\{i||x_i| > r\}$ est fini et a donc sa mesure nulle, par conséquent $\mu(\{i||x_i| \leq r\}) = 1$, d'où ${}^*|a|^* \leq {}^*r$. □

Par exemple $< \frac{1}{i+1} >$ et $< \frac{1}{2^i} >$ sont des infiniments petits > 0.

5.3 Le théorème de Loś

On note Σ la signature correspondant à celle des corps ordonnés (dont $\mathbb{R}$ fait partie) avec un symbole de constante par réel : dit autrement, Σ contient le symbole de relation binaire $\leq$, le symbole de relation binaire $=$, les symboles de fonctions binaires $+$, $\times$, les constantes 0 et 1, et un symbole de constante $\underline{r}$ pour chaque réel r autre que 0 et 1.

13. Autrement dit c'est l'ensemble quotient.
14. Ne dépend pas des représentants de chaque classe.

On note par abus de notation aussi $\mathbb{R}$ (resp. ${}^*\mathbb{R}$) la structure correspondant à $\mathbb{R}$, c'est-à-dire le modèle de cette signature où chaque symbole $\underline{r}$ est interprété par le réel r (resp. ${}^*r$), et où les autres symboles sont interprétés comme dans les réels (resp. comme ci-dessus).

Etant donnée une formule ϕ, on note $\phi(x_1, \ldots, x_n)$ pour indiquer qu'elle possède les variables libres $x_1, \ldots, x_n$. On note $\mathbb{R} \models \phi[a_1, \ldots, a_n]$ (respectivement : ${}^*\mathbb{R} \models \phi[< a_i^1 >, \ldots, < a_i^n >]$) pour signifier que la formule ϕ est satisfaite lorsque $x_1, \cdots, x_n$ prennent les valeurs $a_1, \ldots, a_n$ (respectivement : $< a_i^1 >, \ldots, < a_i^n >$).

Question 15. *Le théorème de Loś s'énonce de la façon suivante :*
Pour toute formule $\phi(x_1, \cdots, x_n)$ sur la signature Σ, pour tous $< a_i^1 >, \ldots, < a_i^n > \in {}^*\mathbb{R}$,

$$ {}^*\mathbb{R} \models \phi[< a_i^1 >, \ldots, < a_i^n >] \text{ si et seulement si } \mu(\{i | \mathbb{R} \models \phi[a_i^1, \ldots, a_i^n]\}) = 1. $$

Sa preuve s'effectue par induction sur la formule ϕ.
Démontrer le cas (significatif) suivant de cette induction :
— le cas où ϕ est la formule $\neg\psi(x_1, \ldots, x_n)$.

Solution :

$$\begin{array}{rcl} {}^*\mathbb{R} \models \phi[< a_i^1 >, \ldots, < a_i^n >] & \Leftrightarrow & \text{non } \psi[< a_i^1 >, \ldots, < a_i^n >] \\ & \Leftrightarrow & \text{non } \mu(\{i | \mathbb{R} \models \psi[< a_i^1 >, \ldots, < a_i^n >]\}) = 1 \\ & \Leftrightarrow & \mu(\{i | \mathbb{R} \models \psi[< a_i^1 >, \ldots, < a_i^n >]\}) = 0 \\ & \Leftrightarrow & \mu(\mathbb{N} - \{i | \mathbb{R} \models \psi[< a_i^1 >, \ldots, < a_i^n >]\}) = 1 \\ & \Leftrightarrow & \mu(\{i | \mathbb{R} \models \phi[< a_i^1 >, \ldots, < a_i^n >]\}) = 1 \end{array}$$

(la première ligne par définition de $\neg\psi$, la seconde par hypothèse d'induction, la troisième par le fait que μ faut 0 ou 1, la quatrième par la propriété 4. ($\mu(X) = 0$ si et seulement si $\mu(\mathbb{N} - X) = 1$), la cinquième par la définition de $\neg\psi$ en chaque entier i). □

On admet tous les autres cas de l'induction (qui ne sont pas plus difficiles), et donc le thèorème.

Question 16. *Utiliser le Théorème de Loś pour montrer que ${}^*\mathbb{R}$ est un corps ordonné (on pourra utiliser le fait qu'être un corps ordonné s'exprime par des formules closes, sans lister ces formules).*

Solution : Etre un corps ordonné s'exprime par un ensemble de formules [15] $\mathcal{T}$. Soit ϕ une formule de ϕ. On a ${}^*\mathbb{R} \models \phi[< a_i^1 >, \ldots, < a_i^n >]$ puisque $\mu(\{i | \mathbb{R} \models \phi[a_i^1, \ldots, a_i^n]\}) = \mu(\mathbb{N}) = 1$. Donc ${}^*\mathbb{R}$ satisfait toutes les formules $\mathcal{T}$ et donc est un corps ordonné. □

Par ailleurs, une conséquence immédiate est un **principe de transfert** : pour toute formule $\phi(x_1, \ldots, x_n)$ de variables libres $x_1, \ldots, x_n$ sur la signature Σ, et pour toutes valeurs $a_1, \ldots, a_n$ dans $\mathbb{R}$,

$$ \mathbb{R} \models \phi[a_1, \ldots, a_n] \text{ si et seulement si } {}^*\mathbb{R} \models \phi[{}^*a_1, \ldots, {}^*a_n]. \quad (1) $$

En particulier, les formules closes satisfaites sur $\mathbb{R}$ et ${}^*\mathbb{R}$ sont exactement les mêmes.

Une fonction $f : \mathbb{R} \to \mathbb{R}$ s'étend en une fonction ${}^*f : {}^*\mathbb{R} \to {}^*\mathbb{R}$ toujours avec le même principe : pour chaque $< x_i > \in {}^*\mathbb{R}$, on pose ${}^*f(< x_i >) = < f(y_i) >$.

15. closes, i.e. n vaut 0 dans ce qui suit.

5.4 Et pourquoi ne pas aller plus loin ...

En réalité, la preuve précédente du théorème de Loś et du principe de transfert qui en découle fonctionne pour la signature Σ, mais fonctionnerait exactement de la même façon pour des signatures étendues et étendant les opérations selon le même principe.

En particulier, pourquoi ne pas le faire dans un langage encore plus riche de telle sorte à permettre de faire référence à tout ce que l'on a envie de manipuler lorsqu'on fait de l'analyse, du moment qu'on exprime des propriétés qui s'écrivent en logique du premier ordre sur cette signature.

Formellement, on considère Σ comme la signature précédente à laquelle on a ajouté des symboles relationnels bien choisi. On obtient alors : toute formule [16] $\phi(x_1, \ldots, x_n)$ qui s'écrit sur la signature Σ satisfait le principe de transfert (c'est-à-dire la propriété (1) pour tout $a_1, \ldots, a_n \in \mathbb{R}$).

Une façon de "bien" choisir les symboles relationnels qu'on ajoute à la signature précédente est d'ajouter un symbole relationnel $\underline{f}$ pour chaque fonction f sur $\mathbb{R}$, de telle sorte que $\underline{f}(x, y)$ code $f(x) = y$. L'intérêt est simplement dans ce qui suit qu'en faisant ainsi on garantit que la formule (2) (comme (3)) dans ce qui suit, est bien (équivalente à) une formule (du premier ordre) sur Σ, et qu'on a bien le principe de transfert.

On va chercher maintenant à utiliser ce résultat.

Par exemple, puisque

$$\forall x \forall y \ \sin(x + y) = \sin(x) \times \cos(y) + \cos(x) \times \sin(y) \tag{2}$$

est vérifiée dans $\mathbb{R}$ (et pourrait bien s'exprimer comme une formule sur Σ) elle doit l'être dans ${}^*\mathbb{R}$.

Dit autrement : puisque la formule est vérifiée dans $\mathbb{R}$ car elle signifie

$$\forall x \in \mathbb{R} \forall y \in \mathbb{R} \ \sin(x + y) = \sin(x) \times \cos(y) + \cos(x) \times \sin(y)$$

on a aussi

$$\forall x \in {}^*\mathbb{R} \ \forall y \in {}^*\mathbb{R} \ {}^*\sin(x + y) \ = \ {}^*\sin(x) \ {}^*\times \ {}^*\cos(y) \ {}^*+ \ {}^*\cos(x) \ {}^*\times \ {}^*\sin(y)$$

vérifiée dans ${}^*\mathbb{R}$.

Autre exemple : puisque e est l'unique racine sur $\mathbb{R}$ de $\ln(e) = 1$, on peut écrire la formule $\phi(e)$ (de variable libre e) définie par

$$\ln(e) = 1 \wedge \forall x \ x \neq e \Rightarrow \ln(x) \neq 1 \tag{3}$$

satisfaite sur $\mathbb{R}$. Cette formule, se transfère à ${}^*\mathbb{R}$, et s'interprète sur ${}^*\mathbb{R}$ comme ${}^*\ln({}^*e) = {}^*1 \wedge \forall x \in {}^*\mathbb{R} \ x \neq {}^*e \Rightarrow {}^*\ln(x) \neq {}^*1$. On en déduit donc que l'unique hyperréel e racine de ${}^*\ln(e) = {}^*1$ sur ${}^*\mathbb{R}$ est le réel ${}^*e$.

Le but des questions qui suivent est de montrer qu'une fonction $f : D \subset \mathbb{R} \to \mathbb{R}$ est continue en un réel r si et seulement si pour tout x infiniment proche, $f(x)$ est infiniment proche de $f(r)$.

Soient $a, b \in {}^*\mathbb{R}$. On note $a \approx b$ (et on dit que a et b sont infiniment proches) lorsque $a - b$ est infiniment petit.

Question 17. *Utiliser le principe de transfert pour démontrer le résultat suivant :*

Soit une fonction $f : \mathbb{R} \to \mathbb{R}$. Soit $r \in \mathbb{R}$ un réel.

$(*)$ *Démontrer que si pour tout $x \in {}^*\mathbb{R}$ tel que $x \approx {}^*r$, on a ${}^*f(x) \approx {}^*f({}^*r)$, alors f est continue en le réel r.*

16. du premier ordre, i.e. de la logique vue en cours

Solution : Supposons que pour tout $x \in {}^*\mathbb{R}$, $x \approx r$, on a ${}^*f(x) \approx {}^*f({}^*r)$.

Soit $\epsilon > 0$.

Soit η un infiment petit > 0, on a

$$\forall u({}^*|u^* - {}^*r|^* \leq \eta \wedge u \neq {}^*r) \Rightarrow {}^*|{}^*f(u)^* - {}^*f({}^*r)|^* \leq {}^*\epsilon.$$

Donc,

$$\exists \eta \in {}^*\mathbb{R}(\eta^* > 0 \wedge \forall u({}^*|u^* - {}^*r|^* \leq \eta \wedge u \neq {}^*r) \Rightarrow {}^*|{}^*f(u)^* - {}^*f({}^*r)| \leq {}^*\epsilon.$$

En transférant cette proposition, on obtient que

$$\exists \eta \in \mathbb{R}(\eta > 0 \wedge \forall u(|u - r| \leq \eta \wedge u \neq r) \Rightarrow |f(u) - f(r)| \leq \epsilon.$$

Puisque c'est vrai pour tout $\epsilon > 0$, f est continue en r. □

Question 18. $(*)$ *Démontrer la réciproque.*

Solution :

Supposons que f est continue en r. Soit $\epsilon \in \mathbb{R}$, $\epsilon > 0$ quelconque. Il y a un réel $\eta > 0$ tel que

$$\forall u, (|u - r| \leq \eta \wedge u \neq r) \Rightarrow |f(u) - f(r)| \leq \epsilon.$$

On transfère cette proposition.

$$\forall u({}^*|u^* - {}^*r|^* \leq {}^*\eta \wedge u \neq {}^*r) \Rightarrow {}^*|{}^*f(u)^* - {}^*f({}^*r)|^* \leq {}^*\epsilon.$$

Soit $x \in {}^*\mathbb{R}$, $x \approx {}^*r$. Puisque $x - {}^*r$ est un infiniment petit, ${}^*|u^* - {}^*r|^* \leq {}^*\eta$ et donc ${}^*|{}^*f(u)^* - {}^*f({}^*r)| \leq {}^*\epsilon$.

Puisque c'est vrai pour tout réel $\epsilon > 0$ réel, ${}^*f(x) \approx {}^*f({}^*r)$. □

Note bibliographique

La partie sur l'analyse non-standard est inspirée du texte "Balade en Analyse non standard sur les traces de A. Robinson" de André Pétry.