corrigé

Bord maximal d'un mot

Question 1.

a. Sachant que $|\beta(v)| \leqslant |v| - 1$, il existe un rang $k \leqslant |v|$ à partir duquel la suite $(\beta^i(v))_{i \geqslant 1}$ stationne en ε. Puisque $\beta^{i+1}(v)$ est préfixe et suffixe de $\beta^i(v)$ il résulte par transitivité que tous les mots de $\{\beta(v), \dots, \beta^k(v)\}$ sont des bords de v.

Réciproquement, si w est un bord de v, notons i l'unique entier vérifiant : $|\beta^i(v)| \leqslant |w| < |\beta^{i-1}(v)|$. Alors w est un bord de $\beta^{i-1}(v)$ donc $|w| \leqslant |\beta(\beta^{i-1}(v))| = |\beta^i(v)|$, ce qui prouve que $w = \beta^i(v)$.

b. Un bord de va est soit égal à ε soit s'écrit wa où w est un bord de v. D'après la question précédente il appartient donc à l'ensemble $\{\varepsilon, \beta(v)a, \dots, \beta^k(v)a\}$.

Tous les mots de cet ensemble sont suffixes propres de va mais pas nécessairement préfixes ; le plus long d'entre eux qui soit préfixe de v est donc le bord de va.

c. On a $j_1 \leqslant k-1$ et $j_{i+1} \leqslant j_i - 1$ donc $j_i \leqslant k - i$ et en particulier $j_{k+1} \leqslant -1$. Mais b est à valeurs dans $\mathbb{N} \cup \{-1\}$ donc $j_{k+1} = -1$.

Posons $v = u_0 \cdots u_{k-1}$ et $a = u_k$. Par définition de $b(k)$ nous avons $\beta(v) = u_0 \cdots u_{j_1 - 1}$ et plus généralement : $\beta^i(v) = u_0 \cdots u_{j_i - 1}$. Ainsi, $\beta^i(v)a$ est préfixe de v si et seulement si $u_{j_i} = u_k$.

Par définition de α, $\beta(v)a, \dots, \beta^{\alpha-1}(v)a$ ne sont donc pas préfixes de v.

Si $u_{j_\alpha} = u_k$, $\beta^\alpha(v)a$ est préfixe de v et d'après la question précédente, $\beta(va) = \beta^\alpha(v)a$, ce qui prouve que $b(k+1) = j_\alpha + 1$.

Si $j_\alpha = -1$ c'est que $\beta^\alpha(v) = \varepsilon$. Dans ce cas, $\beta(va) = \varepsilon$ et $b(k+1) = 0 = j_\alpha + 1$.

Question 2.

```
let bord u =
  let n = string_length u in
  let b = make_vect (n+1) (-1) in
  let rec aux k = function
    | j when j = -1 || u.[j] = u.[k] -> j+1
    | j -> aux k b.(j)
  in for k = 0 to n - 1 do b.(k+1) <- aux k b.(k) done ;
  b ;;
```

Question 3. m est facteur de s si et seulement si m est un bord d'un des préfixes du mot mxs. Puisque x n'est présent ni dans m ni dans s ce bord est maximal. Ainsi, m est facteur de s si et seulement si $|m|$ est présent dans le tableau b.

```
let kmp m s =
  let n = string_length m in
  let b = bord (m ^ "@" ^ s) in
  let rec aux k = b.(k) = n || aux (k+1) in
  try aux 0 with Invalid_argument "vect_item" -> false ;;
```

Le coût de cette fonction (tant temporel que spatial) est un $O(|m| + |s|)$.

Question 4. u et v sont conjugués si et seulement si $|u| = |v|$ et si u est facteur de vv, d'où la fonction :

```
let conjugue u v =
  string_length u = string_length v && kmp u (v ^ v) ;;
```

Question 5. Un mot u contient un facteur carré si et seulement s'il existe un conjugué de u qui possède un bord non vide. En effet, si $u = v_1 w^2 v_2 = v_1 w \cdot w v_2$ alors w est un bord de $wv_2 \cdot v_1 w$. D'où la fonction :

```
let carre u =
  let n = string_length u in
  let rec aux k =
    let b = bord ((sub_string u k (n-k)) ^ (sub_string u 0 k)) in
    b.(n) <> 0 || aux (k+1)
  in try aux 1 with Invalid_argument "sub_string" -> false ;;
```

Question 6. Supposons que v soit période de u ; il existe $n \geqslant 1$ tel que u soit préfixe de v^n. Ceci implique que w est préfixe de v^{n-1} et donc que wv est préfixe de v^n. Puisque $|u| = |wv|$ on en déduit que $u = wv$. Le mot w est donc un bord de u.

Réciproquement, si w est un bord de u, il existe v' tel que $u = vw = wv'$. Considérons un entier n tel que $|v^n| \geqslant |u|$. On a $v^n w = v^{n-1} wv' = \cdots = vwv'^{n-1} = uv'^{n-1}$. Par choix de n il en résulte que u est préfixe de v^n et donc que v est période de u. D'où la fonction :

```
let periode u =
  let n = string_length u in
  let b = bord u in
  sub_string u 0 (n - b.(n)) ;;
```

Question 7. On considère une lettre $x \in \mathrm{A}$ qui n'est pas lettre du mot u et on considère le mot $m = ux\overline{u}$, où $\overline{u}$ est l'image miroir de u. Si v est préfixe de u alors $u = vw$ et $m = vwx\overline{w}\,\overline{v}$ donc v est un palindrome si et seulement si v est un bord de m.

La fonction **bord** permet de calculer le bord maximal $\beta(m)$ de m et donc de déterminer le plus grand des préfixes de u qui soit un palindrome. Les autres préfixes palindromes sont les autres bords de m à savoir $\beta^2(m), \beta^3(m), \cdots$ que l'on détermine à l'aide du tableau des bords.

```
let miroir u =
  let n = string_length u in
  let v = create_string n in
  for k = 0 to n-1 do v.[k] <- u.[n-1-k] done ;
  v ;;

let rec prefixe_palindrome u =
  let n = string_length u in
  let b = bord (u ^ "@" ^ (miroir u)) in
  let rec aux k = match b.(k) with
    | 0 -> []
    | j -> (sub_string u 0 j)::(aux j)
  in aux (2*n+1) ;;
```