

feuille de td 5

1 Exercice *Ponctuation*

Lire et commenter le texte extrait du lexique joint (à considérer comme du cours)

2 Exercice *Marque de formatage*

Un changement de ligne peut suivre un signe de ponctuation mais ne doit pas l'isoler du mot précédent. Selon l'usage, tout signe de ponctuation est donc suivi d'un espace (sauf en fin de paragraphe ou devant un autre caractère d'organisation) et attaché au mot qui le précède :

- Les signes simples (. ,) sont collés au mot précédent.
- Les signes composés (; : ! ?) sont rattachés au mot précédent par un espace insécable (assimilable à une lettre).

Les signes d'encadrement (parenthèses, guillemets, accolades, crochets, etc.) sont collés au texte qu'ils encadrent. À part en tête de paragraphe, les signes ouvrants sont précédés d'un espace. Les signes fermants sont suivis d'un espace, d'une ponctuation ou d'un autre caractère d'organisation.

Les signes documentaires (les puces, tirets, points d'énumération, etc.) peuvent être rattachés au mot suivant par un espace insécable (ou constituer des attributs de paragraphes).

Dans cet exercice, on représente l'espace par •, l'espace insécable par ◦, le tiret ordinaire par -, le tiret insécable par —, le tiret conditionnel par ⇢, rien par ×. l'encadré suivant représente la succession des caractères tapés pour composer un paragraphe. Les passages à la lignes ne coïncident pas avec ceux que produirait le traitement de textes.

Après•une ◻ légère•régression ◻ (◻ de• ◻ 1 ◻ , ◻ 3 ◻ %•en•1 ◻ 993 ◻) ◻ , ◻
les•bénéfices•de•la•S.P.R.•attei ◻ gnaient•3 ◻ 787 ◻ 682 ◻ , ◻ 00 ◻ F•en•1 ◻ 994 ◻ . ◻ C ◻ ,
◻ est•un•record ◻ ! ◻ Le ◻ 2 ◻ avril•dernier ◻ , ◻ 8 ◻ 200 ◻ t•de•NPK•étaient•déjà•vendues ◻ :
◻ 1 ◻ 995•sera•peut ◻ être•encore•meil ◻ leure ◻ . ◻ ¶

Q.1.

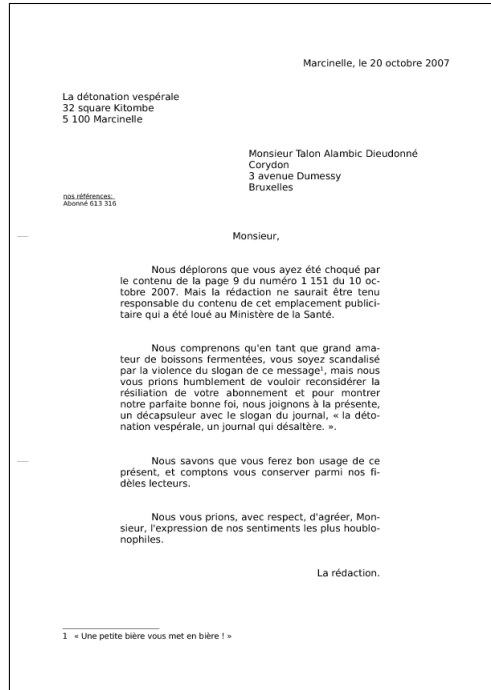
3 Exercice *Attribut de présentation*

Compléter le tableau suivant en indiquant pour chaque attribut s'il concerne, les caractères, les paragraphes, ou les pages.

attribut	concerne
espacement avant le paragraphe 0,5cm	
soulignement en pointillé	
retrait de première ligne -0,5cm	
lignes solidaires	
position en indice 58% points	
taille 16 points	
taquet décimal à 11cm	
Police Bitstream Sans	
Orientation de la page	

4 Exercice à l'ambe¹ –hic– d'yeux donnés

Voici une reproduction d'une lettre reçue par Monsieur Talon.



Sachant qu'on n'a jamais utilisé d'espacement après les paragraphes, Combien y a t'il de paragraphes ? dire quels sont les attributs des différents paragraphes.

5 Exercice *Les expressions régulières*

Une expression régulière est une suite de caractères permettant de représenter un ensemble de chaînes de caractères. Par exemple la chaîne :

`^.*@etudiant.univ-lille1.fr$`

décrit toutes les chaînes se terminant pas `@etudiant.univ-lille1.fr`

Certains caractères, ont un rôle particulier :

- `.` désigne n'importe quel caractère.
- `*` désigne une répétition un nombre quelconque de fois de ce qui précède
- `+` désigne une répétition au moins une fois de ce qui précède.
- `[]` permettent de donner une liste de caractères autorisés.
- `^` placé comme premier caractère signifie le début de la chaîne
- `$` placé comme dernier caractère signifie la fin de la chaîne
- `|` permet de faire une alternative
- `()` permettent de resoudre les problèmes de priorité
- `?` désigne que l'élément qui précède peut être présent zero ou une fois
- `{ }` permettent de préciser le nombre de fois que doit être répéter l'élément qui précède.
- `\` permet lorsqu'il est placé devant un caractère spécial d'annuler l'effet de ce caractère spécial.

Par exemple :

- `^(lundi|mardi|mercredi|jeudi|vendredi)$`
correspond à l'une des 5 chaînes lundi...vendredi.
- `^[0-9a-f][0-9a-f]$`
correspond à une chaîne de deux caractères choisis parmi les chiffres hexadécimaux
- `^([1-9][0-9]{0,3}|0)$`
correspond à une chaîne de 1 à 4 caractères, composée de chiffre ne commençant pas par 0 ou bien à 0 tout seul.
- `^$`
correspond à la chaîne vide.
- `^a*$`
correspond à la chaîne vide, ou bien à a ou bien à aa, ou bien à aaa ...
- `^\.$`

¹Combinaison de deux numéros joués et sortis ensemble

correspond à un . et non pas à n'importe quel caractère.

Q.2. donner une expression régulière correspondant à une date abrégée, c'est à dire par exemple 09/10/2007

Q.3. Donner une expression régulière correspondant à une heure précisée sous la forme 23 h 02 min 05 s

Q.4. Donner une expression régulière correspondant à un verbe du premier groupe. (un verbe qui se termine par er)

Lorsqu'on omet le caractère ^ s'est comme si on autorisait n'importe quoi au début c'est à dire que l'absence de ^ signifie ^. * C'est la même idée pour l'omission du \$.

Q.5. Donner deux expressions rationnelles qui signifient qu'une chaîne doit contenir le mot : **etudiant**

6 Exercice *Germinal*

Voici un extrait du fichier `germinal.txt`

```
Dans_la__plaine_rase,__sous_la_nuit__sans_étoiles,__d'une__obscurité_et¶
d'une__épaisseur_d'encre,__un_homme__suivait_seul__la_grande__route_de¶
Marchiennes_à__Montsou_dix__kilomètres_de_pavé__coupant_tout__droit,__à¶
travers_les_champs__de_betteraves.__Devant_lui,__il__ne_voyait_même_pas¶
le_sol_noir,__et_il_n'avait__la_sensation_de_l'immense_horizon_plat_que¶
par_les__souffles_du_vent__de_mars,__des__rafales_larges_comme__sur_une¶
mer,__glacées__d'avoir_balayé_des_lieues__de_marais_et__de_terres_nues.¶
Aucune_ombre_d'arbre_ne_tachait_le__ciel,__le_pavé_se_déroulait_avec_la¶
rectitude_d'une_jetée,__au_milieu_de_l'embrun_aveuglant_des_ténèbres.¶
¶
L'homme_était_parti_de__Marchiennes_vers_deux_heures.__Il_marchait_d'un¶
```

Le fichier `germinal.txt` contient des espaces surnuméraires, des marques de fin de paragraphes superflues, des lignes vides. Sachant que le texte complet de *germinal* prend plus de 300 pages, il est déraisonnable de faire les modifications à la main, comment s'y prendre ?
