



NOM -----

PRENOM -----

CLASSE -----

- 3) Proposer deux scénarios de tests d'acceptation Given – When –Then (**avec deux CONTEXTES différents**) pour une de ces 'User Story' :

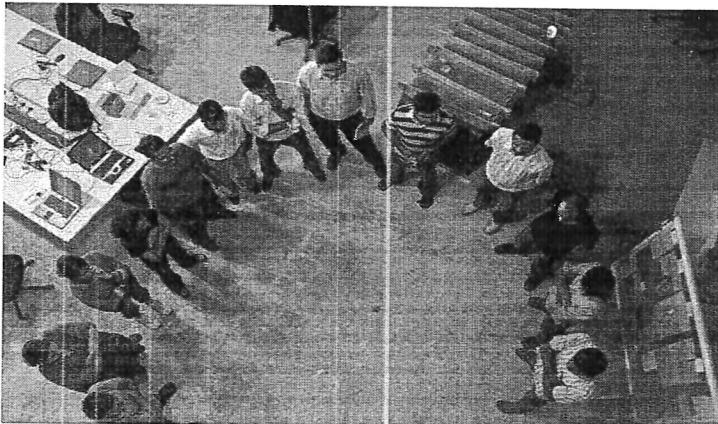
ID de l'user story choisie :

TA1 :- Étant donné que (contexte1)-----

TA2 : Étant donné que (contexte2)-----

BON TRAVAIL 😊

La photo suivante représente un Daily Meeting. Donner 2 Critiques



Critique 1-----

Critique 2-----

EXERCICE 4 **5 Pts**

- 1) Qu'est-ce qu'une 'User Story' ?

- 2) Soit le « feature » : Annuler l'achat d'un article acheté sur le site

Proposer deux User Story associées à ce feature

feature	ID	User Story
Annuler l'achat d'un article acheté sur le site	1	<i>En tant que client, je veux annuler l'achat d'un article que j'ai déjà commandé.</i>
	2	<i>En tant que ----- ----- -----</i>
	3	<i>En tant que ----- ----- -----</i>

EXAMEN

Matière: METHODOLOGIE DE CONCEPTION OBJET

Documents: Non autorisés

Enseignante: N.Daghfous

Durée: 1h30'

Classes : DSI 3*

Prénom----- NOM-----
Classe-----

QCM

5.5 Pts

Lequel de ces principes est mis en avant par le Manifeste Agile ?

1. Valoriser les individus et les interactions, davantage qu'une application qui fonctionne
2. Valoriser les individus et les interactions, davantage que la négociation de contrat
3. Valoriser les individus et les interactions, davantage que les processus et les outils
4. Valoriser les individus et les interactions, davantage que la collaboration avec le client

Dans une équipe agile, qui est responsable de la qualité ?

1. Les développeurs
2. Les testeurs
3. Le product owner
4. Toute l'équipe

Quel est le but de la mêlée quotidienne, ou stand-up meeting ?

1. Le fait de se lever de son siège et de remettre son corps en activité est un team-building en soi.
2. Cela permet, pour chaque membre de l'équipe, de communiquer sur ses avancements et ses points de blocage.
3. Cela oblige l'équipe à travailler en une même unité de lieu.
4. Cela aide à savoir où concentrer ses efforts pour ne pas livrer le projet en retard.

Un client considère qu'une fonctionnalité n'a pas été correctement implémentée. Comment l'équipe agile doit-elle réagir ?

1. En ajoutant une tâche d'analyse de ce problème au backlog de l'équipe
2. En corrigeant ce point, toutes affaires cessantes
3. En rédigeant un rapport d'anomalie, qui sera traité une fois que le backlog sera vide
4. En modifiant les spécifications et en demandant au client de les valider avant d'effectuer les modifications dans l'application

Lors d'une rétrospective, qui peut légitimement émettre des remarques sur l'activités de test ?

1. Les développeurs seulement
2. Les testeurs seulement
3. Tout le monde sauf les testeurs
4. Toute l'équipe

Soit la user story suivante : 'En tant que propriétaire de chien, je souhaite enregistrer mon chien en ligne afin d'être rapidement en règle avec les lois en vigueur.' Pour que cette user story soit en accord avec les critères INVEST, quelle est la dimension à développer ?

1. Sécurité
2. Valeur
3. Testabilité
4. Les trois

NOM _____
 PRENOM _____
 CLASSE _____

Que mesure un burn-down chart ?

1. Le temps prévu en fonction du budget
2. La qualité
3. La rapidité d'exécution du travail
4. La quantité de travail qui reste à faire en fonction du temps restant

Lequel de ces savoir-faire ou savoir-être est le plus utile pour un testeur en contexte agile qu'en contexte de projet séquentiel ?

1. La collaboration
2. Le test manuel
3. Le test de performance
4. La gestion des cas de test

Lors de la première séance de planning de poker, les points de story suivants ont été donnés à une US en fonction du risque, de l'effort, de la complexité et de l'étendue des tests.

5
5
20

Quel est le meilleur résultat suivant cette séance de planning Poker ?

1. On choisit 5 qui est répété 2 fois
2. On discute et un autre tour planning de poker devrait se produire après la discussion
3. Le scrum master décide
4. Le product owner décide

Quel est l'intérêt de se restreindre à seulement quelques améliorations par itération ?

1. Si l'on planifie trop de choses, l'équipe risque de se décourager et de perdre sa motivation.
2. Si l'équipe essaie d'implémenter trop de choses à la fois, cela peut perturber son rythme de livraison.
3. Si, lors d'une itération, l'équipe livre trop de modifications, le client risque d'en demander autant à chaque fois.
4. Aucun intérêt.

Que peut-on raisonnablement attendre d'un cas de test au cours du temps ?

1. Qu'il ne change pas
2. Qu'il soit automatisé et qu'il ne demande qu'une maintenance minime
3. Qu'il évolue au rythme du logiciel qu'il concerne
4. Qu'il soit édité par le développeur qui implémente la fonctionnalité associée

EXERCICE 2

4.5 Pts

Les méthodes agiles s'appuient sur trois piliers. Citer ces piliers en précisant un moyen permettant d'assurer chacun de ces piliers avec brève explication

Pilier 1 : _____

Moyen d'assurer le Pilier 1 : _____

Comment ce moyen assure le Pilier 1 : _____

Pilier 2 : _____

Moyen d'assurer le Pilier 2 : _____

Comment ce moyen assure le Pilier 2 : _____

Pilier 3 : _____

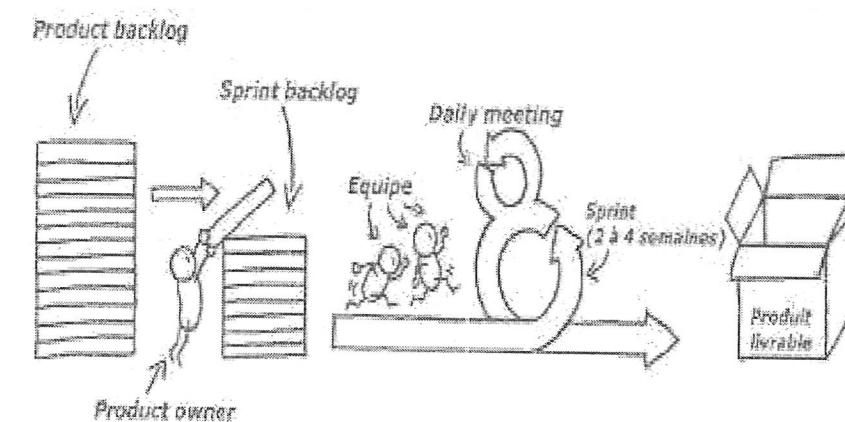
Moyen d'assurer le Pilier 3 : _____

Comment ce moyen assure le Pilier 3 : _____

EXERCICE 3

4 Pts

Le schéma suivant représentant un Sprint Scrum. Donner 2 Critiques pour cette schématisation.



Critique 1 : _____

Critique 2 : _____

EXERCICE 2 MAPREDUCE**(4 POINTS)**

On souhaite établir des statistiques sur la longueur des mots dans un document à partir d'une application *Map-Reduce*.

Appliquer le principe MAP Redue pour implémenter un programme qui à partir d'un document (ensemble de mots), il peut générer le nombre de mots proportionnellement à la longueur de mots (nombre de caractères)

Sachant qu'on a en entrée un fichier contenant la phrase suivante :

« *Un sourire coûte moins cher que l'électricité, mais donne autant de lumière* »

En sortie on devra avoir :

- Le nombre de Mots composé de 1 à 4 caractères (inclus)
- Le nombre de Mots composé de 5 à 8 caractères (inclus)
- Le nombre de Mots composé de plus de 8 caractères

Question 1 : Appliquer le principe du MapReduce sur cette phrase en entrée en effectuant 4 étapes (**faites un schéma explicatif**) :

1. *Distribuer les données sur au moins 3 partitions qui doivent être relativement équilibrés*
2. *Appliquer le programme du Map Clé valeur en précisant quelle est la clé du programme Map*
3. *Shuffle and sort la sortie du programme Map*
4. *Donner le résultat final de Reduce après l'exécution de différentes étapes*

EXAMEN

Matière: GESTION DE DONNÉES MASSIVES

Documents : Non autorisés

Enseignants: Mhafdh. H, Ghorbel. I, Limam.I

Durée: 1h30

Classes : DSI3 x

Date : Jan 2021

Le devoir comporte 4 pages. Il vous est demandé d'apporter un soin particulier à la présentation de votre copie.

EXERCICE1 : QCM**(8 POINTS)**

Pour chaque question choisir une ou plusieurs réponses correctes

- 2) Le principe de réPLICATION est conçu pour garantir :
 - a- La scalabilité
 - b- La cohérence
 - c- La tolérance aux pannes
 - d- Le parallélisme
- 3) Le principe de PARTITIONNEMENT est conçu pour garantir :
 - a- La disponibilité
 - e- La scalabilité
 - f- La tolérance aux pannes
 - g- La cohérence
 - h- Acidité
- 4) Comment fonctionne la distribution de fichiers sur HDFS ?
 - a- Répartition en fonction de la taille des fichiers sur chaque nœuds du cluster.
 - b- Répartition en blocs répliqués sur les nœuds du cluster.
 - c- Répartition en nœuds répliqués sur les blocs du cluster.
 - d- Répartition en fonction des choix de l'utilisateur au moment de l'upload
- 5) Quel est le rôle du NameNode ?
 - a- Ecrire Écrire ou lire les données sur les DataNodes.
 - b- Vérifier la disponibilité des données sur les DataNodes.
 - c- Remplacer un DataNode si un d'entre eux devient indisponible.
 - d- Administre les transactions en autorisant ou non la lecture / écriture des fichiers.
- 6) CASSANDRA est une base de données :
 - a- Hiérarchique
 - b- Orienté document
 - c- Orienté graphe
 - d- Relationnelle
 - e- Orientée colonne

Bon travail

7) Les colonnes en HBASE sont organisées sous forme de

- a- Column group
- b- Column families
- c- Column list
- d- Column

- e- Connector API
- f- Topic API

8) Le langage de requête de Hive est:

- a- Hive latin
- b- Hive QL
- c- HiveSQL
- d- hiveShell
- e- HQL

9) YARN est l'acronyme de

- a- Yet Another Reserve Negotiator
- b- Yet Another Resource Network
- c- Yet Another Resource Negotiator
- d- Aucune de ces réponses

10) Vous pouvez exécuter Pig en mode batch en utilisant

- a- Pig shell command
- b- Pig scripts
- c- Pig options

11) Lequel des éléments suivants a occupé la place de TaskTracker de Yarn

- a- NodeManager
- b- ResourceManager
- c- ApplicationMaster
- d- Scheduler

12) Laquelle des affirmations suivantes est correct pour RDD dans SPARK?

- a- RDD est un paradigme de programmation
- b- RDD dans Apache Spark est une collection immuable d'objets
- c- RDD est une base de données

13) Apache Kafka possède 4 principales APIs lequels:

- a- Producer API
- b- Consumer API
- c- Brokers API
- d- Streams API

EXERCICE 2 NOSQL

(4 POINTS)

Soit une base de données relationnelle de vente de livres comportant un ensemble de tables dont un extrait est décrit ci-dessous avec 3 tables (Vente, Livre et Auteur).

TABLE VENTE

ticket	date	books
1	01/01/2016	2154889522
1	01/01/2016	2154889589
2	01/01/2016	2154889589

TABLE LIVRE

books	titre	id_auteur
2154889522	Asterix et Cléopatre	154
2154889589	NoSQL	987

TABLE AUTEUR

id_auteur	Nom	prénom
154	Gosciny	René
987	Bruchez	Rudi

On souhaite faire la migration de cette base à une base de données NoSQL orientée document.

Pour cela Proposer pour la table vente, une modélisation orientée document. (C'est-à-dire représenter les données de la table vente sous forme de document)

Question 1 :

- Qu'est-ce qu'une base de données NoSQL ?
- Quelles sont les propriétés qui caractérisent une base NoSQL ?

Question2 : proposez deux exemples de document pour représenter les données décrits dans les tables ci-dessus

Clé1	{ } }
Clé2	{ } }

EXAMEN DE SYNTHESE

Matière: ECOSYSTEME BIGDATA AVANCEE

Documents: Non autorisés

Enseignant : Inès Ghorbel

Durée: 1.30h

Classe : RSI31

Date: 29/01/2021

QCM

2/ Laquelle des réponses suivantes est considérée comme le principal avantage de Spark par rapport à MapReduce ?

- Spark supporte Scala.
- Spark peut exécuter des calculs en mémoire.
- Spark supporte les RDD.
- MapReduce supporte le calcul distribué.

3/ Comment fonctionne la distribution de fichiers sur HDFS ?

- Répartition en fonction de la taille des fichiers sur chaque nœuds du cluster.
- Répartition en blocs répliqués sur les nœuds du cluster.
- Répartition en nœuds répliqués sur les blocs du cluster.
- Répartition en fonction des choix de l'utilisateur au moment de l'upload

4/ Les lacunes de Hadoop MapReduce ont été surmontées par Spark RDD à travers :

- Lazy-evaluation (Évaluation paresseuse)
- DAG
- In-memory processing (Traitement en mémoire)
- Tout ce qui précède

5/ L'API dataset est accessible dans

- Java et Scala
- Java, Scala et python
- Scala et Python
- Scala et R

6/ Lequel des énoncés suivants n'est pas une transformation ?

- Flatmap
- Map
- Reduce
- Filter

7/ Laquelle des réponses suivantes est considérée comme le principal avantage de Spark par rapport à MapReduce ?

- Spark supporte Scala.
- Spark peut exécuter des calculs en mémoire.
- Spark supporte les RDD.
- Map Reduce supporte le calcul distribué.

8/ Quel langage n'est pas supporté par spark

- Java
- Python
- Scala
- Ruby

9/ Quel est le rôle du NameNode ?

- Écrire ou lire les données sur les DataNodes.
- Vérifier la disponibilité des données sur les DataNodes.
- Remplacer un DataNode si un d'entre eux devient indisponible.
- Administrer les transactions en autorisant ou non la lecture / écriture des fichiers.

10/ Quel démon est responsable de la réPLICATION de données dans Hadoop?

- HDFS.
- Task Tracker.
- Job Tracker.
- Name Node.
- Data Node

Questions de cours

Spark

1/ Que retourne chacune de ses transformations ou action de rdd :

a : scala> val longueursLignes = texteLicence.map(l =>l.length)

scala>longueursLignes.take(10)

b : scala>lignesAvecCopyright.take(4)

c : scala>texteLicence.first()

d :scala>texteLicence.take(10)

e : scala>texteLicence.collect()

2/ L'architecture succède rapidement et s'interface avec des solutions de type/mapreduce pour effectuer des calculs, à haute performance, pour des données de grande ou très grande taille. Spark, qui est initialement développé en, fournit une interface de programmation riche et en développement.

3/ Quels types d'Opérations est supporté par un RDD, donner un exemple par opération ?

4/ Avec quels 'composants' SPARK se positionne dans l'écosystème bigdata

Batch/ETL processing:.....

Stream processing:.....

Machine Learning:.....

Interrogation SQL-Like:.....

Graph processing:.....

YARN

1/ Présentez l'architecture de YARN dans l'écosystème BigData (les différents composants de YARN).

2/ Présentez dans cette architecture le déroulement et les différentes étapes de traitement et d'analyses avec spark des fichiers log et elastick search et de visualisation des résultats (hadoop, spark , cassandra ELK , Zookeeper , ...) sachant que il existe des données des anciennes analyses dans la base de données nosql Cassandra.

3/ Quel est l'avantage de yarn dans hadoop

Etude de cas :

Le **Tracking** a ici sa place sur l'ensemble de la chaîne, de la conception du produit jusqu'à son arrivée chez le client de UPS. **Avec la démocratisation des technologies de suivi et d'identification**, la masse de données disponibles dans ce secteur va exploser.

Récemment **UPS** (UPS : leader mondiale dans le service logistique comme Aramex en Tunisie), en utilisant les données collectées via les GPS de l'ensemble de ses véhicules, a revu l'ensemble de ses trajets pour la troisième fois de son histoire afin d'économiser de l'essence, et du temps.

Celui-ci calcule en temps réel le meilleur trajet pour le véhicule en fonction du contenu de son chargement. Cela évite au chauffeur de saisir une adresse et lui permet d'éviter l'ensemble des embouteillages.

Cela a également permis à UPS de développer un nouveau service pour ses clients, MyChoice, qui leur permet de modifier l'heure et le lieu de livraison jusqu'à la dernière minute. L'entreprise gagne donc en flexibilité tout en offrant un service unique à ses clients.

Une autre technologie, ILC, permet de surveiller les conditions environnementales auxquelles sont soumis les produits tout au long de la chaîne logistique. **Cette technologie qui produit énormément de données** ne peut être analysée que via le Big data **afin de les traiter en temps réel** et signaler tout dysfonctionnement ou anomalie aux opérateurs afin qu'ils puissent déclencher des opérations de maintenance plus rapidement ou détecter le vol de marchandises avec un système d'authentification à travers le site web ou l'application Mobile.

1/ Proposez une architecture bigdata pour UPS, qui va subvenir à tous les nouveaux services que UPS souhaite intégrer ?

Bon travail

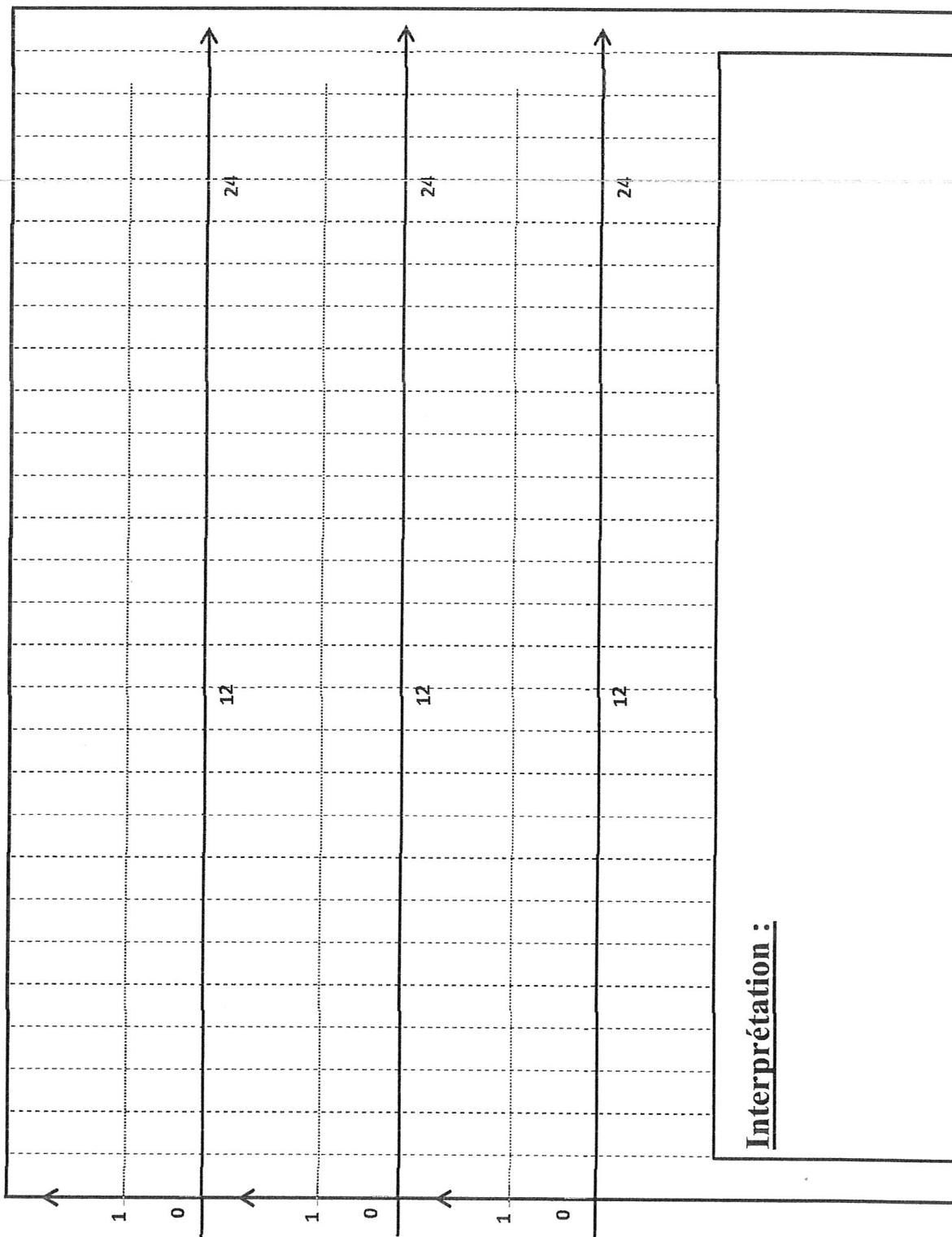
NOM et Prénom :

CLASSE :

NOM :
Prénom :

CLASSE :

DR5



DR1

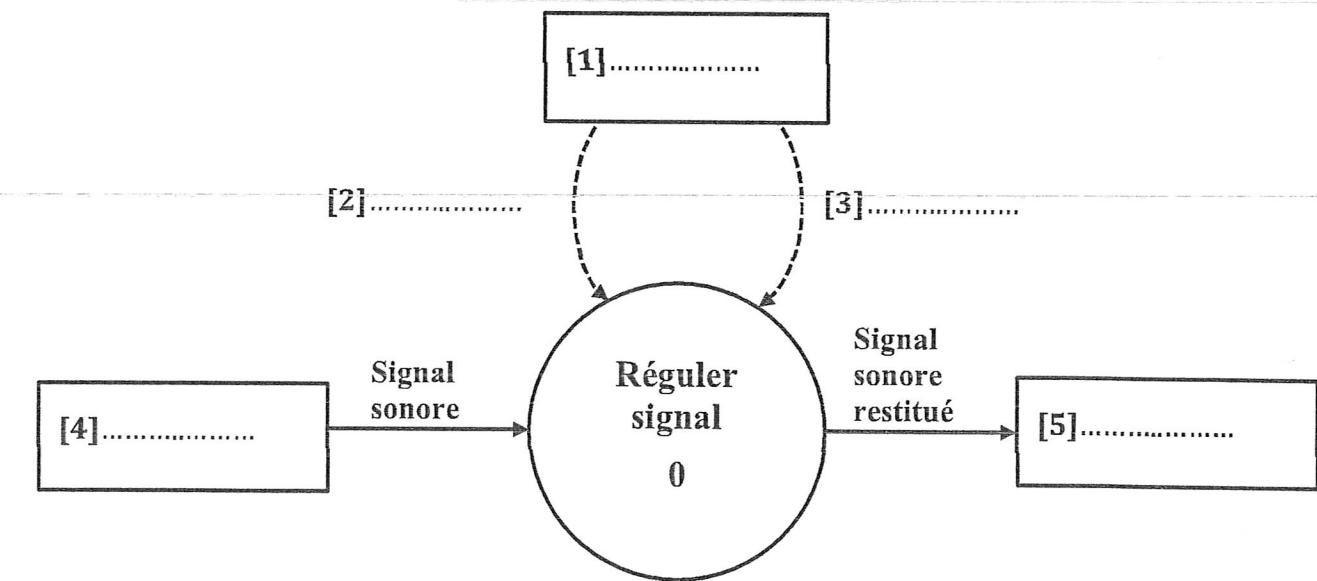


Figure (2) : Diagramme de contexte

DR2

[1] :	[2] :
[3] :	[4] :
[5] :	[6] :
[7] :	[8] :
[9] :	

NOM et Prénom :

CLASSE :

NOM et Prénom :

CLASSE :

DR3

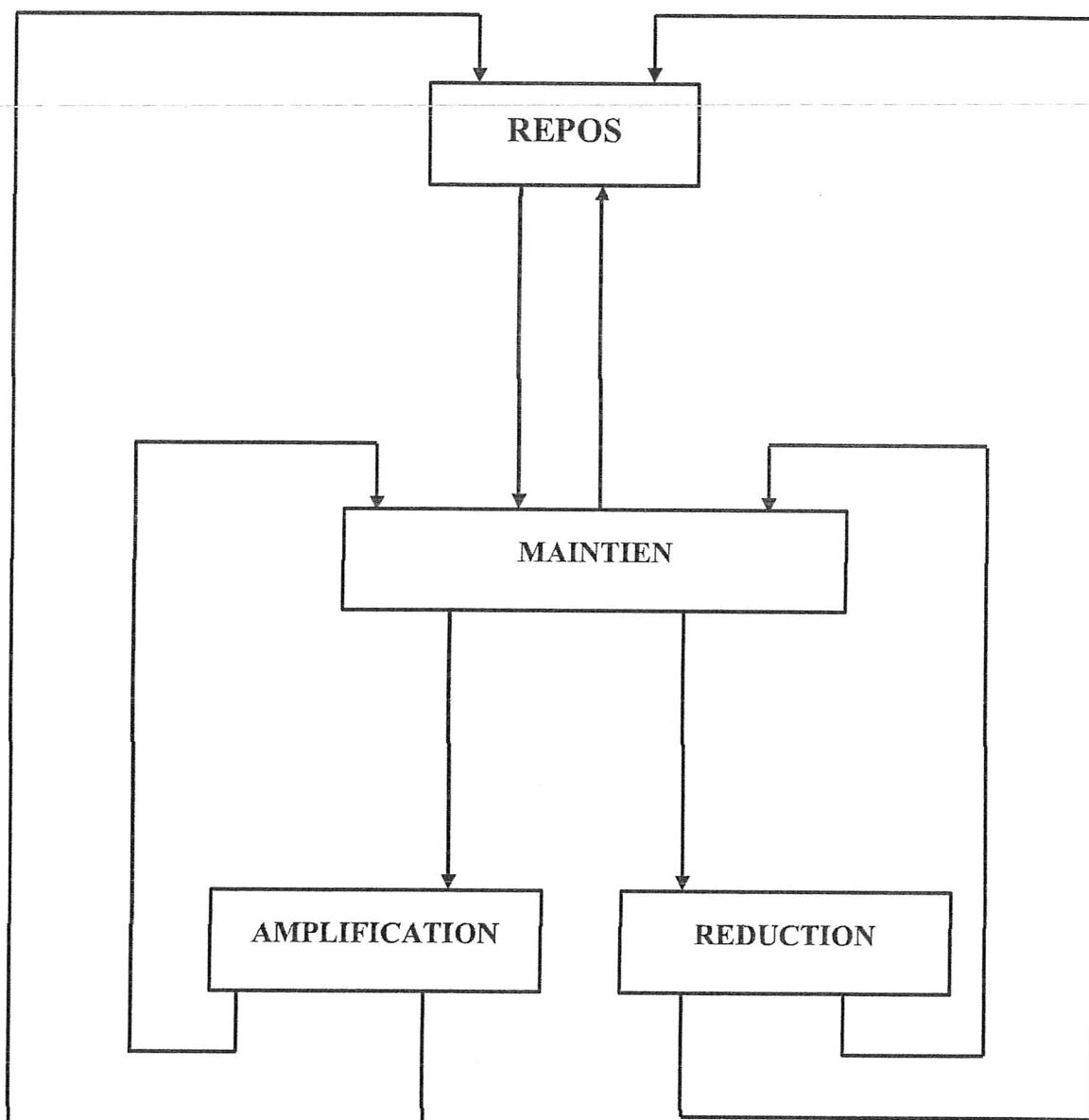
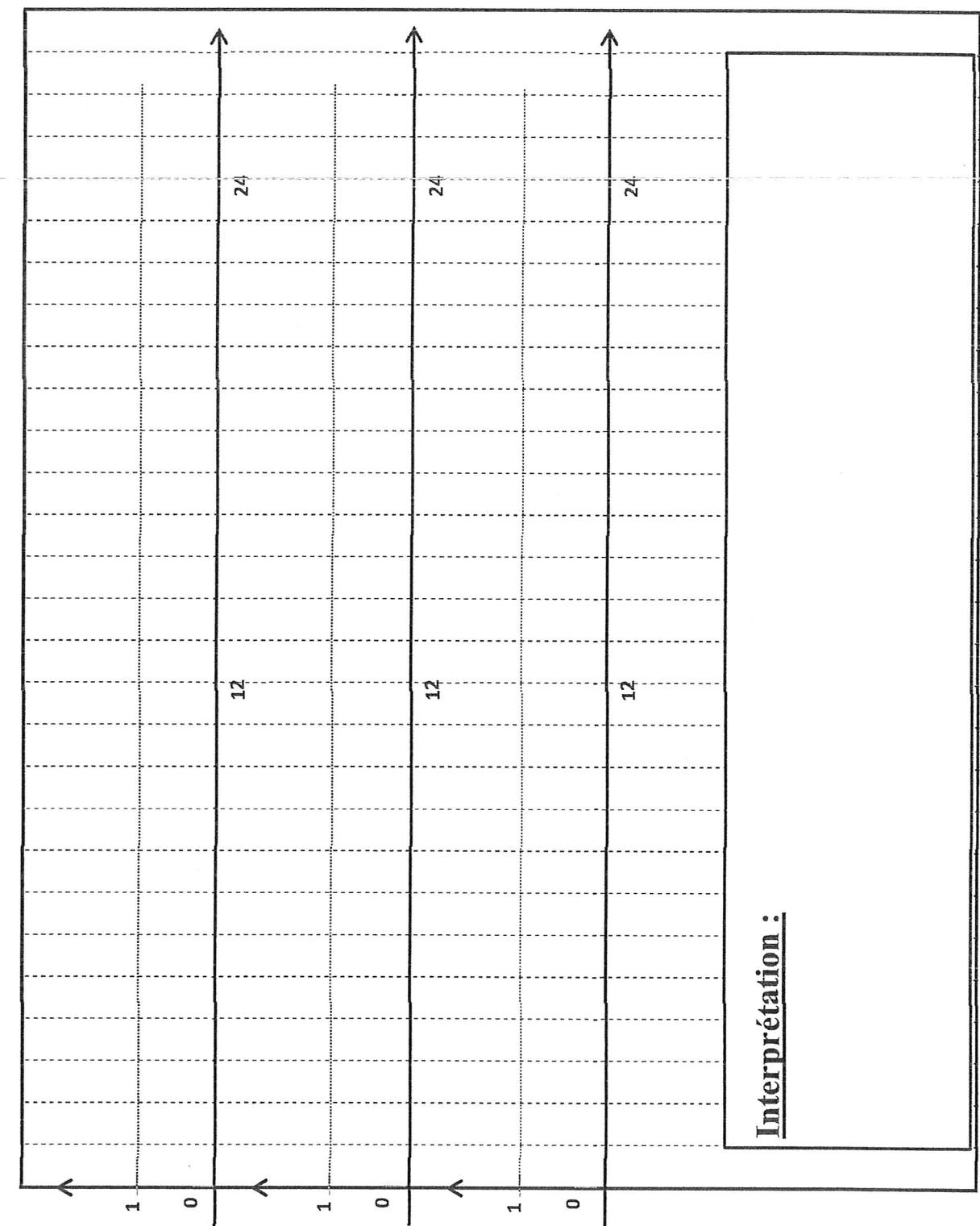


Figure (4) : Diagramme Etat/Transition

DR4



Département	Génie électrique	Epreuve	Examen
Spécialité	Automatismes et Informatique Industrielle	Date	Janvier 2020
Classes	AII31, AII32	Durée	1.5 Heures
Matière	Systèmes temps réel	Documents	Non autorisés
		Calculatrice	Autorisée

Le sujet comporte 2 pages d'énoncé et 4 pages de document-réponse.

Il sera tenu compte de la présentation des copies et, de la clarté des réponses.

Exercice (1) (11 points)

Dans cet exercice on se propose de contrôler l'amplitude d'un signal sonore à travers une application temps réel. Le système est composé d'un **microphone** qui permet de capter le signal sonore. L'application permet de saisir et de stocker le signal sonore. Un processus fonctionnel de « détection de seuil » vérifie l'amplitude du **signal saisi** par rapport à un seuil maximal et un seuil minimal. En fonction de l'état du signal l'application **amplifie** ou **réduit** l'amplitude de ce signal.

Le signal restitué est transmis vers un **haut-parleur**. Une console de commande permet la mise en **marche** et l'**arrêt** du système de régulation.

1°/ Compléter le diagramme de contexte du système (DR1) (2.5 points)

2°/ Compéter le tableau (DR2) relatif au diagramme préliminaire présenté dans la figure (1). (4.5 points)

3°/ Déterminer le diagramme d'Etat /Transition du processus de contrôle en respectant la structure donnée par le schéma dans le document réponse (DR3) (4 points)

Exercice (2) (9 points)

On considère les tâches $T_i = (C_i, P_i)$ suivantes, définies par leur capacité et leur période : $\{(3,8), (5,12), (4,24)\}$

1°/ Dressez le diagramme d'exécution des trois tâches selon l'algorithme EDF sur le document-réponse DR4. (3 points)

2°/ Conclure et indiquer le nombre de préemption (1 point)

3°/ Dressez le diagramme d'exécution des trois tâches selon l'algorithme RMA sur le document-réponse DR5. (3 points)

4°/ Conclure et indiquer le nombre de préemption. Quel est le meilleur algorithme ? Expliquez (2 points)

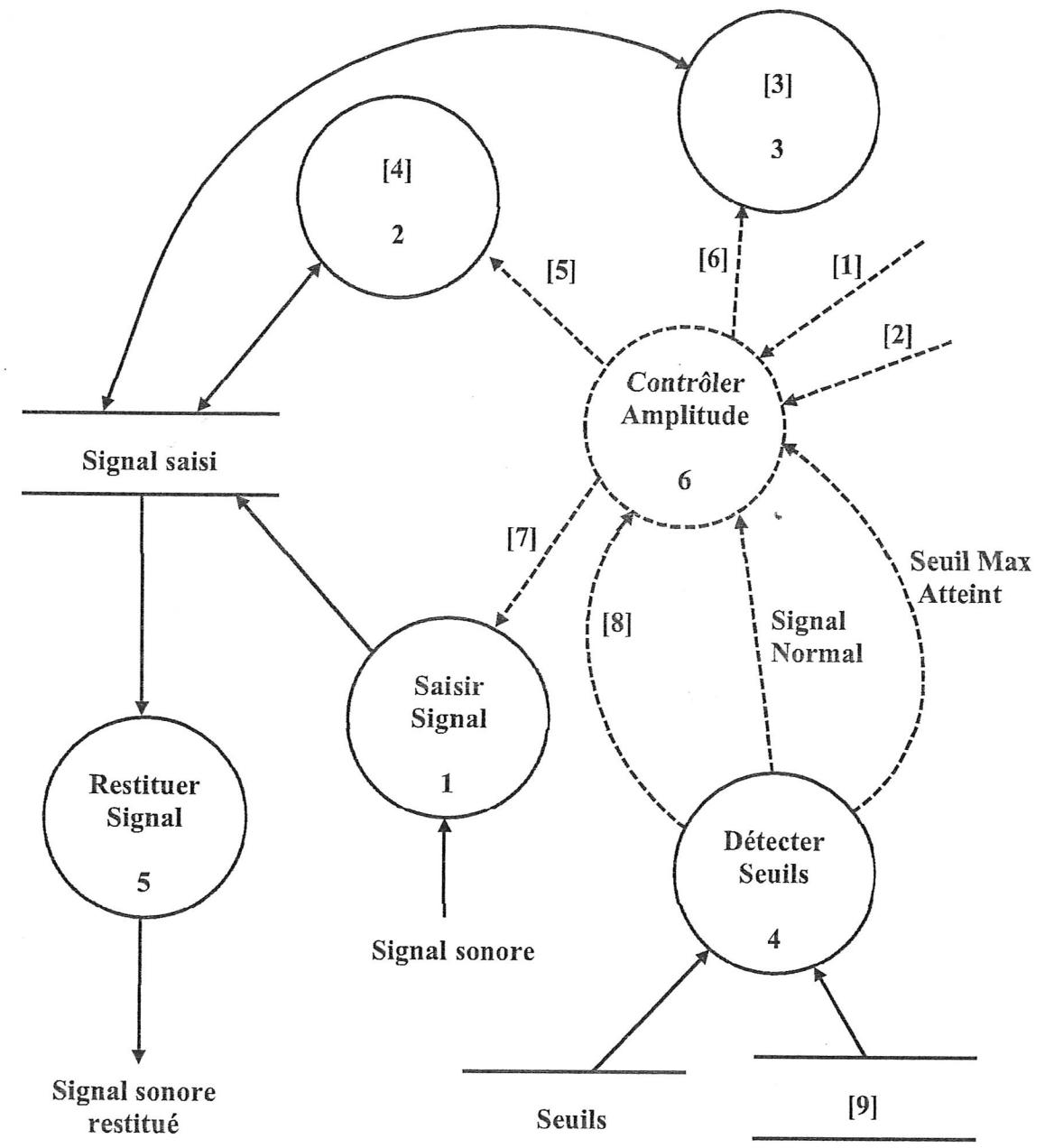


Figure (1) : Diagramme préliminaire