

Protocole expérimental

07/01/2016

Introduction

L'outil que vous allez utiliser est un prototype expérimental permettant de tester une approche théorique : permettre d'exprimer des processus d'analyse de traces d'apprentissage indépendamment de toute plate-forme d'analyse.

Ce prototype n'étant pas toujours intuitif, la première partie va permettre de vous familiariser avec l'outil. La deuxième partie va vous proposer d'exprimer par vous-même un processus d'analyse spécifique. Dans les deux cas, vous n'aurez qu'à décrire le processus d'analyse, sans procéder à son implémentation dans une plate-forme d'analyse de traces spécifique.

Il est à noter que l'outil est encore en phase de développement, que des bugs subsistent encore, et que certaines fonctionnalités (comme des opérateurs) ne sont pas encore implémentées.

Étape préalable

Objectif

Avant de commencer à utiliser l'outil, cette étape préalable consiste à exprimer un processus d'analyse sur papier, comme vous le feriez en temps normal pour réfléchir au problème posé par un demandeur.

Cette étape a pour but d'étudier la manière dont les analystes et les statisticiens conceptualisent le besoin d'un demandeur, avant d'entamer la récolte des données et leur manipulation pour arriver au but escompté par ledit demandeur.

Nous vous proposons donc de décrire sur papier un processus d'analyse simple, de votre choix, que nous vous demanderons de reproduire un peu plus tard dans l'outil.

Si vous n'avez pas d'idée, nous vous proposons de travailler sur le processus d'analyse qui porte sur un indicateur : **“Pourcentage, pour la semaine dernière, de visionnage d'une vidéo donnée”**.

La page suivante vous permet de décrire votre processus.

Description du processus

Objectif du processus :

Complexité (subjectif, selon vous) du processus :

Description des étapes du processus (textuelle, graphique ou les deux) :

Se familiariser avec l’outil

Description de l’outil

L’outil que vous allez utiliser tente d’offrir aux analystes une méthode pour exprimer les processus d’analyses sans avoir à se soucier des différentes plates-formes d’analyses à utiliser.

Pour ce faire, vous ne manipulerez pas directement les données aux travers des opérateurs comme vous pouvez en avoir l’habitude, mais plutôt le type de ces données. Par exemple, vous indiquerez qu’il faut un nom, une date, un score... mais jamais “Dupont”. De cette manière, vous fixez des pré-requis sur le contenu de vos données.

Pour décrire votre processus, vous utiliserez des opérateurs, qui prennent en entrée certains types de données et qui en produisent d’autres en sortie.

Il est important que vous compreniez que la partie de l’outil que vous allez utiliser est uniquement descriptif et n’a pas vocation à calculer : elle ne sert qu’à exprimer un processus. L’instanciation - ou l’application de la description du processus à un vrai cas d’étude - pour indiquer/effectuer les démarches nécessaires et réaliser le calcul est envisagé dans une autre partie qui sera développée à terme.

Cas 1 : Prise en main (durée 30 min.)

Dans cet exemple, qui va vous aider à comprendre la logique de fonctionnement de l’outil, nous allons essayer de représenter le processus d’analyse de l’indicateur **“nombre de personnes connectées à une plate-forme dans un intervalle de temps précis”**.

Nous allons procéder étape par étape, mais avant cela, il faut réfléchir à comment obtenir notre processus.

Pour l’obtenir, on peut procéder de cette manière : “Il faut d’abord isoler l’intervalle de temps. Puis, compter les apprenants qui ont passé le filtre”.

1. Présentation de l’outil

Chargez la page “create_gap.html” présente sur le support USB ou à l’adresse http://elearning-dev.univ-lyon1.fr/IOGAP/create_gap.html?. Vous devez avoir accès à internet pour que l’outil fonctionne.

C’est dans cette interface que vous évoluerez. Considérez-la comme le “cahier de brouillon” dont vous vous servirez pour exprimer votre réflexion. La 1^{ère} zone grise, en dessous du titre “Créateur de Processus d’Analyse Générique”, représente les ressources à votre disposition (listes initiales, opérateurs, listes calculées). La 2^{nde} zone grise représente votre processus d’analyse en cours de construction (partie gauche), ainsi que sa représentation graphique (partie droite).

La dernière zone, cachée par le bouton “Sélection des étapes”, permet de sélectionner les étapes qui constitueront au final le processus d’analyse.

2. Initialisation du processus

Dans la 2nde zone, cliquez sur l'onglet "+" afin d'indiquer au système que vous souhaitez commencer à créer un processus d'analyse. Vous verrez apparaître un nouvel onglet "Feuille 1" contenant 3 colonnes (Inputs, Operator, Outputs).

3. Définition des données relatives aux traces d'apprentissage

La première chose à faire consiste à définir les *données nécessaires* (ou plutôt le type des données) que vous utiliserez. Cliquez sur le bouton "**Listes Initiales**". Vous verrez apparaître une nouvelle fenêtre.

Les listes initiales représentent les données nécessaires dont vous avez besoin pour effectuer votre processus d'analyse. Elles peuvent être utilisées à n'importe quelle étape du processus.

Cliquez sur le bouton "**Créer une nouvelle liste**", et remplissez les champs requis avec les informations relatives à vos types de données. Dans notre exemple, nous pouvons ainsi créer une liste "MOOC_avec_video". N'oubliez pas de valider. La liste apparaît alors avec 0 élément.

Une donnée nécessaire est toujours contenue dans une liste. Cette étape de création d'au moins une liste est obligatoire.
Une liste représente ce que doit impérativement contenir votre trace.

Cliquez sur la liste qui vient d'apparaître. Vous pouvez maintenant ajouter des éléments à cette liste. Il faut en ajouter deux (le temps et les apprenants). Cliquez sur "**Ajouter un élément**" pour ajouter le premier élément. Recommencez l'opération pour le deuxième.

En cliquant deux fois sur "Ajouter un élément", vous spécifiez donc deux données nécessaires impératives à votre processus d'analyse.
Notez que lors de leur création, les nouvelles données ont un nom attribué par défaut du style "autoElementX".

Pour sélectionner une donnée nécessaire, il suffit de cliquer dessus.

Le fait de sélectionner un élément permet d'appliquer les actions "Supprimer l'élément", "Utiliser l'élément" et "Modifier l'élément".

Modifiez le nom des deux données créées. Appelez par exemple l'une "temps", l'autre "idApprenant".

La zone de modification ne s'ouvre qu'après sélection d'un élément. Notez que renommer

convenablement vos éléments permet de mieux vous organiser. C'est la sémantique que donne l'utilisateur aux données nécessaires qu'il utilise qui apporte un sens au processus.

L'attribut *type* n'a pas encore d'impact, vous pouvez laisser *xs:type*.

4. Ajouter un opérateur

La première étape de notre processus d'analyse étant un filtre temporel sur des données temporelles, il faut ajouter "temps" à l'étape courante via "**Utiliser l'élément**" (bouton vert). Vous pouvez fermer la fenêtre.

Si vous ajoutez malencontreusement plus d'un élément, vous pouvez retirer ceux en trop en fermant la fenêtre, en sélectionnant -pour chaque élément indésirable- l'élément à supprimer puis en cliquant sur "Supprimer l'élément sélectionné".

Par ailleurs, quand vous fermez la fenêtre des listes initiales, vos données sont conservées et pourront être réutilisées pour un autre opérateur. Pas d'inquiétude, donc.

Ensuite, il nous faut matérialiser l'action que l'on va appliquer sur ce type de données. Il s'agit ici d'un filtre temporel. Cliquez sur le bouton "**Opérateurs**". Une nouvelle fenêtre s'ouvre.

"Opérateurs" recense toutes les actions (opérations) que notre outil est capable de gérer (actuellement) et qu'il sera capable d'instancier dans les plates-formes connues d'analyses de traces.

Cliquez sur l'opérateur "Filtre temporel" pour le sélectionner.

Lorsque vous sélectionnez un opérateur dans la liste, toutes ses informations sont accessibles, notamment le nombre de données nécessaires qu'il attend en entrée.

Utiliser l'opérateur sélectionné en cliquant sur "**Utiliser cet opérateur**".

Le fait d'avoir suffisamment d'entrées pour un opérateur dans une étape génère automatiquement le résultat.

Le fait de cliquer sur le bouton ajoute directement l'opérateur à l'étape courante. Vous ne pouvez ajouter qu'un seul opérateur à chaque fois.

Si vous désirez remplacer l'opérateur de l'étape courante, il suffira d'utiliser un autre opérateur qui remplacera l'ancien.

Cliquez sur "ok" pour fermer la fenêtre d'alerte, puis cliquez sur le bouton "Fermer". Vous revenez alors sur l'interface principale.

Un opérateur génère toujours une ou plusieurs listes calculées.

Ici, la liste de données calculées contiendra les mêmes éléments que l'ancienne. Ce n'est pas toujours le cas.

Dans notre exemple, nous appliquons un filtre temporel : cette action n'est pas destructrice sur les types de données présents dans les traces, uniquement sur les données qu'ils contiennent.

5. Configurer l'opérateur

Afin de configurer l'opérateur pour cibler la période, cliquez sur le bouton vert de l'opérateur dans l'étape courante. Une nouvelle fenêtre s'ouvre.

En cliquant sur un opérateur dans une étape, vous avez accès à un récapitulatif vous permettant de savoir si oui ou non l'opérateur est bien configuré.

Lorsqu'il manque des informations capitales (ou qu'il est surconfiguré), le bouton "Opérateur" est de couleur orange.

Cliquez sur "**Configuration**".

Il y a deux manières de paramétrer un opérateur. La première est celle que nous allons utiliser. La seconde pourra s'effectuer lors de l'instanciation.

Mettons la date de cette semaine, le format étant en JJ/MM/AAAA. Le paramètre 0 sera donc à JJ-7/MM/AAAA et le paramètre 1 à JJ/MM/AAAA. N'oubliez pas de valider les changements.

Dans la prochaine version du prototype, les paramètres et leurs noms seront plus explicites.

Les paramètres vont être utilisés lors de l'instanciation. Ils n'ont pas d'impact dans la zone de création du processus.

Remarquez qu'il est possible de changer le nom de la liste produite en double-cliquant sur la liste dans la colonne Outputs.

Vous venez de finir la première étape de votre processus d'analyse.

6. Ajouter un autre opérateur

Maintenant, place à la deuxième et dernière étape. Cliquez sur “+”, à côté de “Feuille 1”.

Vous pouvez créer autant d'étapes que vous le voulez. Celles incomplètes sont indiquées en orange. Notez que vous pouvez également les supprimer.

Vous pouvez renommer les étapes *via* le bouton “Renommer l'étape courante”.

Le second opérateur de notre processus d'analyse devra compter le nombre d'apprenants restant dans le créneau temporel défini. Pour ce faire, il ne faut pas utiliser “idApprenant” de la liste initiale, mais celui de la liste générée à l'étape 1.

Cliquez sur “**Listes calculées**”.

Si vous aviez utilisé l’“idApprenant” de votre liste initiale, vous auriez compté tous les apprenants, qu'importe la date associée. Là, vous utilisez les “idApprenant” résultant du filtre temporel.

Une fenêtre s'ouvre avec la liste produite à l'étape 1. Un simple clic sur le nom de la liste affiche les informations de cette liste, un double clic fait apparaître les options de manipulation de la liste. Sélectionnez “idApprenant” et utilisez-le dans l'étape 2, puis fermez la fenêtre.

Les listes calculées vous permettent donc de réutiliser les productions des différents opérateurs, et ainsi les enchaîner pour créer votre description du processus d'analyse.

À l'étape N, vous pouvez utiliser toutes les productions des étapes N-1. Vous pouvez créer de cette manière des processus complexes.

Si vous avez ajouté trop de fois un élément, ou que vous l'avez ajouté dans une mauvaise étape, vous pouvez toujours le supprimer.

Il faut maintenant compter les apprenants. Ouvrez le menu “**Opérateurs**” et utilisez “**Compter Distinct**”.

“Compter Distinct” est un opérateur permettant de ne compter qu'une fois le même élément. Ainsi, lors de l'instanciation, si il y a 4 fois l'apprenant avec l'id 777, il sera compté une seule fois.

7. Aperçu graphique du processus

Lorsque vous configurez une étape, le graphe du processus que vous êtes en train de définir se met à jour pour vous permettre d'avoir un meilleur aperçu du processus. Ce graphe est visible sur la partie droite de l'interface. Vous pouvez interagir sur les noeuds du graphe en les sélectionnant.

Dans ce graphe, un cercle est la représentation d'une "feuille" (i.e. étape) du processus d'analyse. Toutes les feuilles, entièrement configurées ou non, seront représentées par des cercles portant leur nom.

Si une feuille utilise des données issues d'une autre feuille, alors cette dépendance est matérialisée par une flèche entre les deux cercles qu'il faut lire "B utilise au moins une donnée de A".

Sur une flèche entre un cercle A et un cercle B apparaît le nom des plate-formes qui supportent l'opérateur défini dans la feuille A.

De plus vous avez la possibilité sous le graphique de choisir une "Visualisation spécifique" en fonction des plate(s)-forme(s) d'analyse de traces supportant les opérations. Les étapes non supportées sont grisées et les étapes non configurées ne sont pas affichées.

8. Finalisation du processus d'analyse

Vous avez terminé d'exprimer toutes les étapes de votre processus d'analyse. Maintenant, il faut choisir lesquelles conserver. Pour ce faire, cliquez sur "Sélection des étapes" en bas de la page et choisissez les deux étapes créées puis sauvegardez.

La possibilité de sélection permet de créer des étapes brouillonnes et de ne pas les prendre en compte quand vient le temps d'exporter le processus d'analyse.

Le système sélectionne automatiquement les étapes dépendantes d'une étape choisie, et ainsi de suite. Aucun risque d'oublier une étape importante de cette manière !

Félicitations, vous venez de créer votre première description de processus d'analyse de manière indépendante des plates-formes d'analyse de traces.

L'étape d'instanciation de ce processus n'est pas implémentée pour l'instant, mais devrait par la suite vous permettre de décrire l'indicateur produit, de lier ce processus avec vos traces réelles et d'avoir la notice permettant de l'instancier sur la ou les plate-forme(s) que vous aurez choisie(s).

Questionnaire à remplir.

Cas 2 : Mise en oeuvre de votre processus d'analyse

Maintenant, vous allez tenter de mettre en oeuvre par vous-même le processus que vous avez décrit dans la première partie de l'expérimentation.

Le but est de matérialiser, dans l'outil, la réflexion que vous avez eue lorsque vous avez créé votre processus d'analyse.

Si vous avez des difficultés pour formaliser votre processus, vous pouvez vous référer à la section ci-dessous. Elle se concentre sur l'exemple du **“Pourcentage, pour la semaine dernière, de visionnage d'une vidéo donnée”** :

Pour obtenir ce processus, on peut procéder de cette manière : “Il faut d'abord isoler la semaine désirée. Puis, parmi tous les apprenants disponibles, filtrer ceux qui ont vu la vidéo, puis les compter. Ensuite, compter le nombre total d'apprenants. Et pour finir diviser le nombre d'apprenants ayant vu la vidéo par le nombre total d'apprenants.”

Pour réaliser ce processus, que l'on peut exprimer en 5 étapes, il faut deux données nécessaires : le “temps” et “hasWatched?”, un booléen indiquant si oui ou non l'utilisateur a vu la vidéo.

Essayez maintenant de créer ce processus. Pour ce faire, rechargez la page “*create_gap.html*” et recommencez les manipulations faites pour le processus précédent.

Questionnaire à remplir.