

# Rapport de projet

## Optimisation stochastique 2017

*Arnaud Steinmetz*

*Yannis Weishaupt*

### Le projet

Le [sujet d'optimisation 2017](#) portait sur la création d'un algorithme à l'aide de l'outil [easea](#) qui est un logiciel open source développé par SONIC ([Stochastic Optimisation and Nature Inspired Computing](#)).

L'algorithme doit permettre d'optimiser les créneaux de personnes effectuant des présentations durant un événement donné avec certaines contraintes décrites dans le sujet.

Le projet se divise en deux parties :

- Un parseur qui extrait les données d'un fichier Excel, les traite et les renvoie dans un fichier.txt en respectant un format spécifique.
- Un programme easea appliquant un algorithme d'optimisation.

- 1) Extraction des données : <https://github.com/MrYawe/CS-DC17-parser>
- 2) Programme easea : <https://github.com/MrYawe/CS-DC17>

### Extraction des données

Nous avons choisi de ne pas exploiter les données se trouvant dans l'onglet "All e-tracks" car d'après les échanges avec Paul Bourguine, certaines données étaient manquantes et les papiers ajoutés tardivement n'y figuraient pas. Nous avons donc utilisé les données dans les onglets qui suivent (15, 4, 5, 6, ...).

### Modifications Excel

Nous avons modifiés quelques données dans le xls avant d'y appliquer le parser :  
Le paper n°505 de l'onglet track 6, avait comme valeur 12 dans la colonne track.

→ remplacement de la valeur track de 12 à 6

Le paper n°504 de l'onglet track 12, avait comme valeur 6 dans la colonne track.

→ remplacement de la valeur track de 6 à 12



## Résultat du parser

Le programme écrit en go produit le fichier résultat “*papers.txt*” dont chaque lignes représente un paper sous le format suivant:

PaperID | TrackID | SessionID | Duration | UTC | Constraint 1 | Constraint 2 | Constraint 3

En voici un exemple:

358|4|67|30|2|[6:00,16:00]|[6:00,16:00]|

## Les contraintes de temps

Toutes les contraintes de temps ont été uniformisés, elles sont donc toutes dans le format:

[H:M,H:M]

Avec H l’heure sur un ou deux chiffres et M les minutes sur un ou deux chiffres.

Nous avons dû procéder à quelques modification sur les données d’entrées.

- Si les contraintes de l’excel étaient sous le format[H, H] alors nous les avons transformés sous format [H:00, H:00]
- Si une des deux heures était “24” alors nous l’avons remplacé par “23:59”
- Si avec [H1:M1, H2:M2] on a H2:M2 > H1:M1 alors le programme nous avertit qu’il y a une erreur. Dans ces cas là nous avons corrigé les données à la main dans le xls pour ne plus avoir de warnings.
- Si le programme n’a pas réussi à parser une contrainte nous avons aussi un avertissement.

## La durée des papiers

La durée des papiers n'étant pas défini de base dans l’onglet “all e-tracks” nous avons décidé de la calculer en fonction du type dans le parser.

- 120 minutes pour le type “Tutorial”
- 60 minutes pour le type “Plenary talk”
- 30 minutes pour les “Invited talk”, “Advanced Introduction invited talk”, “New Result invited paper”, “Full paper” et “Young researcher”.
- 15 minutes pour les “Short paper”
- 5 minutes pour les “Poster”



## Le programme easea

Dans le programme nous avons choisi d'utiliser 4 classes à l'extérieur de User Class pour stocker les données.

- Constraint : représentant les contraintes horaire d'une personne
- Paper :
- Session
- Track

Dans un premier temps nous chargeons les données du txt dans une liste de track accessible globalement (vector<Track\*>). Cela nous permet de conserver la hiérarchie du découpage.

Dans la déclaration des User classes nous avons :

- Date : personnalisée pour gérer le temps simplement
- Match : représentant une session
- GenomeClass : qui représente notre génome et possède un tableau de match

Ensuite les 4 étapes importantes :

- Initialisation:

Nous avons donné à chaque session une date de départ aléatoire et un ordre de passage aléatoire pour chaque papier. Pour la date de départ nous nous sommes assuré qu'elle ne dépasse pas la fin de l'événement en prenant en compte la durée d'une session.

- Crossover:

Dans le crossover, nous avons opté pour que le fils hérite tous les attributs d'un des deux parents.

- Mutation

Dans la fonction de mutation, la date de départ et l'ordre de passage peuvent muter indépendamment. La mutation de l'ordre de passage correspond à un nombre aléatoire de permutation. La mutation de la date correspond à 3 mutations indépendante des heures, minutes et secondes définis aléatoirement en faisant attention aux bornes.

- Evaluator :

Dans l'évaluator deux choses étaient essentielles à évaluer.

- Si le Track15 (sessions plénières) est en superposition avec une autre session alors on applique un malus important.



- Les papers doivent avoir une date qui rentre dans le créneaux des personnes qui présentent.

Pour cela, nous avons créé une fonction qui calcule le temps de recouvrement en minute entre deux intervalles de temps.

Pour tout calendrier (génom) qui possède une session/paper) se déroulant en même temps qu'un paper du track15, nous mettons son score à 0. Car l'on ne souhaite aucun recouvrement.

Pour que les papers respectent le créneaux des personnes, nous avons choisi d'ajouter au score le recouvrement en min entre les intervalles données par la personne et la date du paper. Un bonus supplémentaire (multiplication du score du paper par 2) est attribué lorsque celui-ci se retrouve entièrement dans un créneau.

#### Information /!\ :

- **Après discussion par mail avec Paul Bourguine, lorsqu'une colonne de contrainte était vide, alors on suppose que la personne est disponible sur la journée entière.**
- **Les créneaux indiqués par les personnes ont été supposé comme les créneaux où les personnes étaient disponibles.**
- **Les utc ne sont pas pris en compte, mais il serait simple de les utiliser en ajoutant la valeur à la date testé lors de l'évaluation des créneaux.**  
(Nous avons décidé de ne pas le rajouter, car il y avait de nombreux formats différents, et nous ne comprenions pas comment interpréter les valeurs avec et sans la précision utc)
- **Nous avons également remarqué que le track15 est assez particulier, et qu'il y a par exemple un paper Welcome qui est supposé être au début. Nous avons traité les éléments du track15 comme des sessions normales, mais il faudrait plutôt écrire directement ces dates en dur (sinon le welcome peut se retrouver au plein milieu de l'événement)**

## Résultats

Le meilleur résultat (génom) est exporté dans un fichier nommé *events.ics* au format ical, où l'on associe l'id des papers à une intervalle de dates. Il est possible de visualiser





## **Améliorations**

Durant ce projet, nous avons supposé que le track 15 ne pouvait contenir que des sessions plénières. Pour laisser un peu plus de liberté, il aurait été intéressant de mettre un booléen aux sessions dites plénières. Cela permettrait d'avoir des sessions plénières dans d'autres tracks.

Pour une prise en compte directe (sans toucher le code), il serait possible d'appliquer le décalage UTC des personnes directement avec le parseur excel->txt, qui devrait décaler les créneaux de disponibilités en fonction des fuseaux horaires.

Il pourrait également être intéressant d'utiliser un score pour tenter d'éloigner les eSessions les uns des autres, pour permettre d'optimiser l'étalement des sessions sur toute la durée de l'événement.

