

Notes de réunion

WP3

WP3.1

Livrable L3.1.1 (T12) en retard (thèse Jérôme, guidage planifié, ImproteK 4.0) à déposer maintenant (T18), Livrable 3.1.2 (T24) en avance (guidage réactif), devrait être prêt avant Noël (SoMax 2.0). Le WP3.1 aura donc trois mois d'avance.

WP3.2

Adaptation temporelle de l'interaction. Thèse de Kevin Sanlaville, soutenance à mi-parcours décalée à juin 2016, prolongation de six mois (> avr 2017) financée par Labex Smart pour mettre en place étude expérimentale et rédaction. Modèle : Réseaux d'influence entre HMM. Agents GRETA /UNITY 5. L3.2.1 (T18) ne devrait pas poser de problème (2 articles) et L3.2.2 (T30) devrait passer à T24 (mars/avril 2017) donc bien en avance. Explication ANR : l'expérimentation a porté surtout sur les agents conversationnels, et de nouveaux modèles probabilistes issus de WP2 (clusters graphiques) viennent compléter la panoplie de Kevin (HMM).

WP3.3

Agents créatifs

l'intégration logicielle générale du projet débute en fait par ce module dans lequel les approches libres, planifiées et réactives (les trois évolutions du projet OMax) et les modèles d'adaptation temporelles doivent déjà être intégrés. Il est convenu de démarrer par une réflexion partagée sur l'architecture théorique (acquisition, apprentissage, scheduling général, modèles d'interaction, écoute etc. etc.) débouchant sur un document clair, avant les choix de plate-formes, de langages et d'implémentation.

WP2

WP2.1

Ken organise durant son séjour à l'Ircam les expériences d'écoute sur oracle, module probabiliste multi-dimension, et cluster graphique Prob. Avec Axel : comparer cluster graphique et agents adaptatifs soMax sur l'exemple de l'amaïn gauche et la main droite du Köln concert de Keith Jarrett. Les développements nouveaux autour du cluster graphique créent un lien avec WP3.2 Adaptation temporelle de l'interaction (Kevin) dans la mesure où les dimensions qui communiquent par message peuvent devenir des agents autonomes.

WP2.2

Démarrage de la thèse de Nathan Lieberman. Réorientation du WP, moins sélection de dimension, mais deep learning et apprentissage par renforcement. Du coup le WP3.3 apprentissage multi-échelle sera alimenté à la fois par la thèse de Ken (modèles probabilistes) et celle de Nathan (deep learning)

Notes de réunion

WP1

WP1.1

séparation informée multi-canal, travaux d'Antoine, Emmanuel & al sur la réduction d'interférence multi-canal par NMF et la séparation à base de deep learning. Le paquet contient des avancées multiples et variées, et le livrable L1.1 à T+12 sera finalement installé en T+18 (maintenant) Antoine pourrait monter une démo sur les enregistrements du MJF avec suppression d'interférences dans les voix séparées, sur les enregistrements disponibles à l'Inria, ou sur la prochaine campagne (Diane Reeves, xxx).
 Expériences sur le Mixage souhaitées par Antoine (étudiants de la FSMS, superviseur RIM ou ingé-son)

WP1.2 / WP 1.3

Sylvain recrute deux post-doc le plus rapidement possible : un postdoc de 18 mois de T+18 à T + 36 pour WP1.2 écoute structurante, et un post-doc de 12 mois de T+24 à T+36 sur WP1.3 Décomposition / Recomposition.
 Le principe mis à jour de WP1.2 se concentre sur l'idée d'information par les entrées successives d'instruments construisant un passé, et par la structure connue ou inférable (« cristalline » cf modèle anticipation / contraste de F. Bimbot) du Morceau.
 (« extraction polyphonique informée par le passé et connaissance de la structure »)
 Pour WP1.2 comme de WP1.3, il s'agit de décomposition polyphonique (structures de niveau intermédiaire signal/symbolique) permettant de délivrer les flux d'entrée symboliques exploités par WP2 et WP3, plus que de séparation avec reconstitution haute fidélité du signal original.
 Modèle inspiré de Shazam (appariement de pics sinusoidaux dans différentes fenêtres d'analyse), cadre théorique proche de celui de la détection de signaux.

WP4

(expérimentation, collaborations artistiques, dissémination, intégration etc.)

- enregistrements MJF déjà disponibles : multipiste : hancock / corea, BB kNg groupe 12, Huey Lewis., 20 concerts stereo video
- démos à prévoir : nettoyage multi-canal (cf WP1.1)
- résidences musicales : B. Delbecq, G. Nouno, A. Slater, B. Lubat, C. Kely ...
- scénarios musicaux, résidences musicales (T+24-> fin) :
 - enlever une voix du multipiste, nettoyer les interférences de cette voix sur toutes les autres (WP1.1), puis divers scenarios « music minus one » avec instrumentiste live, générations machines, combinaisons des deux (WP2.3). La génération peut être inférée de l'extraction et la modélisation d'un instrument dans un autre document MJF (WP1.1, 1.2, 1.3)
 - Projet Bloch / Sellin / Reeves : travail sur une collection de concepts de Diana Reeves, scenario multiples de décomposition, recomposition, interaction avec le piano live.

Notes de réunion

Langages

existant:

- Inria : Python
- Sylvain : 3 niveaux :
 - interfaces (e.g TCI - TK)
 - structures et processus complexes (Lisp, SBCL, octave)
 - horsepower (signal C, C++)
- Ircam : interaction Lisp (MCL) / Max / ANtescofo (improteK). Interaction Max / Python (SoMax)

Futur :

à déterminer rapidement en fonction d'abord d'une meilleure vision

- du champ applicatif et des utilisateurs du logiciel visé (tous les musiciens ? profils experts RIM / musiciens chercheurs)
- de ce que nous apprend la réflexion sur l'architecture théorique
- des besoins forts en dynamicité (recompilation à la volée, changement du code pendant la performance même etc.) des utilisateurs experts
- de l'évolution du paysage technologique (PyExt pas fiable pour l'avenir ?...)
- des ressources réalistes de maintenance (e.g. un environnement à 3 langages / plate-formes est quasi sûr de s'effondrer une fois que les projets ANR sont finis, que les postdocs attaquent d'autres projets et. etc.)
- des espoirs de diffusion. Idem un environnement à 3 langages/plate-formes est à peyu près sûr de ne pas marcher une fois arrivé chez l'utilisateur qui n'est pas dans le cénacle de l'équipe ou du labo
- etc.

A faire

- inscrire sujet de thèse de Nathan sur le site Dyci2 (Nathan, Emmanuel)
- pdf présentation d'Antoine, pointeurs vers maquette logicielle et exemples audio.
- pour postdoc WP1.2 et 1.3 contacter Nathan X (à Lannion) qui connaît le modèle Shazam, diffuser l'annonce de recrutement sur Atiam Alumni etc. (Sylvain, Emmanuel, Antoine)
- pour expériences mixage Antoine, recrutement ingé-son et proposition de projet à la FSMS par Georges.

Techniques / data set évoqués à regarder :

- adversarial learning (I. Goodfellow)
- database Esling / Leopold, 250 couples midifiles piano reduction / orchestre complet
- Dataset LAKH 170.000 alignement audio - midi

Idées

multipiste Hancock / Corea, tester agents adaptatifs soMax et cluster graphique prob. Ken.