



HAL
open science

Intégration de tâches: étiquetage morpho-syntaxique, analyse syntaxique et analyse sémantique traités comme une tâche unique

Timothée Bernard

► To cite this version:

Timothée Bernard. Intégration de tâches: étiquetage morpho-syntaxique, analyse syntaxique et analyse sémantique traités comme une tâche unique. Traitement Automatique des Langues Naturelles, 2021, Lille, France. pp.268-269. hal-03265885

HAL Id: hal-03265885

<https://hal.science/hal-03265885>

Submitted on 23 Jun 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Intégration de tâches: étiquetage morpho-syntaxique, analyse syntaxique et analyse sémantique traités comme une tâche unique

Timothée Bernard

Laboratoire de linguistique formelle, Université de Paris, France
timothee.bernard@u-paris.fr

RÉSUMÉ

Nous présentons des résumés en français et en anglais de l'article (Bernard, 2021), présenté lors de la conférence *16th Conference of the European Chapter of the Association for Computational Linguistics* (EACL 2021). L'article décrit l'intégration de tâches, un ensemble de principes orthogonaux au partage de paramètres dont le but est de maximiser l'interaction entre différentes tâches. L'intégration de tâches est illustrée avec un système analysant de manière jointe les niveaux morpho-syntaxiques, syntaxiques et sémantiques. La stratégie adoptée par ce système, entraîné par renforcement, est aussi analysée.

ABSTRACT

Multiple Tasks Integration: Tagging, Syntactic and Semantic Parsing as a Single Task

We present abstracts in English and in French for (Bernard, 2021), a paper was presented at the 16th Conference of the European Chapter of the Association for Computational Linguistics (EACL 2021). The paper proposes Multiple Tasks Integration (MTI), a set of principles orthogonal to weight sharing the aim of which is to maximise the interaction between different tasks. MTI is illustrated with a system that performs part-of-speech tagging, syntactic dependency parsing and semantic dependency parsing. The strategy inferred by this system, trained by reinforcement learning, is also analysed.

MOTS-CLÉS : analyse sémantique, analyse syntaxique, étiquetage morpho-syntaxique, apprentissage multi-tâche, apprentissage par renforcement.

KEYWORDS: semantic parsing, syntactic parsing, POS tagging, joint processing, multitask learning, reinforcement learning.

1 Résumé en français

À la recherche d'alternatives autant aux chaînes de traitement séquentielles qu'aux systèmes mono-tâches, nous proposons l'intégration de tâches (MTI, pour *Multiple Tasks Integration*), un paradigme multi-tâche orthogonal au partage de paramètres. L'idée centrale de l'intégration de tâches est de traiter l'entrée simultanément sur différents niveaux d'analyse (i) de manière à ce que chaque décision repose sur l'ensemble des structures alors prédites et (ii) sans imposer aucune des contraintes d'ordre habituelles. De cette manière, les différentes tâches peuvent interagir pleinement les unes avec les autres. En particulier, nous ne forçons pas certains niveaux d'analyse à être analysés avant d'autres, et ni le début de la phrase avant la fin.

Nous illustrons l'intégration de tâches avec un système analysant de manière jointe les niveaux morpho-syntaxiques, syntaxiques et sémantiques. L'intégration de tâches repose sur le calcul, à chaque étape, pour chaque token, d'un score pour chacune des actions possibles (sélections d'une étiquette morpho-syntaxique, sélection d'une tête syntaxique et d'une étiquette de dépendance, etc.) à partir d'un encodage de la phrase et des annotations déjà prédites. C'est à partir de ces scores que sont choisies les actions à appliquer, sans distinctions de type. Nous observons que la mise en place d'un apprentissage par renforcement ainsi que l'abandon des contraintes d'ordre mènent tous deux à un gain de performance sur les tâches d'analyse syntaxique et sémantique. Nous observons aussi que notre modèle adopte une stratégie de type *easy-first*, consistant — en moyenne — à prédire les dépendances en commençant par les plus courtes, mais que le niveau syntaxique n'est pas toujours analysé avant le niveau sémantique.

2 Abstract in English

Departing from both sequential pipelines and monotask systems, we propose Multiple Tasks Integration (MTI), a multitask paradigm orthogonal to weight sharing. The essence of MTI is to process the input iteratively but concurrently at multiple levels of analysis, where each decision is (i) based on all of the structures that are already inferred and (ii) free from usual ordering constraints. This way, the different tasks can fully interact with each other. In particular, we do not constrain the system to perform any given task before any other one, nor to analyse the beginning of the sentence before its end.

We illustrate MTI with a system that performs part-of-speech tagging, syntactic dependency parsing and semantic dependency parsing. The tasks integration is based on a scoring, at each step, for each token, of each possible actions (selection of a POS tag, selection of a syntactic head and a dependency label, etc.) using an encoding of both the sentence and the predictions from previous steps. The actions to perform are selected on the basis of these scores, regardless of their type. We illustrate MTI with a system that performs part-of-speech tagging, syntactic dependency parsing and semantic dependency parsing. We observe that both the use of reinforcement learning and the release from sequential constraints are beneficial to the quality of the syntactic and semantic parses. We also observe that our model adopts an *easy-first* strategy that consists, on average, of predicting shorter dependencies before longer ones, but that syntax is not always tackled before semantics.

Références

BERNARD T. (2021). Multiple Tasks Integration : Tagging, Syntactic and Semantic Parsing as a Single Task. In *Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics : Main Volume*, p. 783–794, Online : Association for Computational Linguistics.