

2.6 Traduction machine

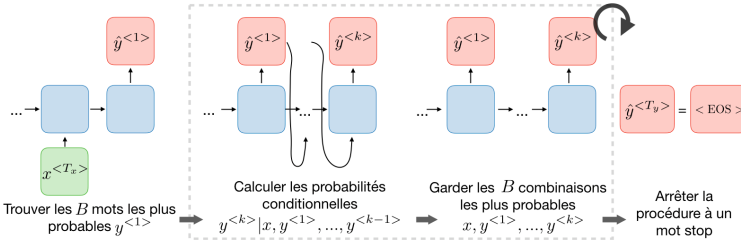
□ **Vue d'ensemble** – Un modèle de traduction machine est similaire à un modèle de langage ayant un auto-encodeur placé en amont. Pour cette raison, ce modèle est souvent surnommé modèle conditionnel de langage.

Le but est de trouver une phrase y telle que :

$$y = \arg \max_{y^{<1>}, \dots, y^{<T_y>}} P(y^{<1>}, \dots, y^{<T_y>} | x)$$

□ **Recherche en faisceau** – Cette technique (en anglais *beam search*) est un algorithme de recherche heuristique, utilisé dans le cadre de la traduction machine et de la reconnaissance vocale, qui vise à trouver la phrase la plus probable y sachant l'entrée x .

- Étape 1 : Trouver les B mots les plus probables $y^{<1>}$
- Étape 2 : Calculer les probabilités conditionnelles $y^{<k>} | x, y^{<1>}, \dots, y^{<k-1>}$
- Étape 3 : Garder les B combinaisons les plus probables $x, y^{<1>}, \dots, y^{<k>}$



Remarque : si la largeur du faisceau est prise égale à 1, alors ceci est équivalent à un algorithme glouton.

□ **Largeur du faisceau** – La largeur du faisceau (en anglais *beam width*) B est un paramètre de la recherche en faisceau. De grandes valeurs de B conduisent à avoir de meilleurs résultats mais avec un coût de mémoire plus lourd et à un temps de calcul plus long. De faibles valeurs de B conduisent à de moins bons résultats mais avec un coût de calcul plus faible. Une valeur de B égale à 10 est standard et est souvent utilisée.

□ **Normalisation de longueur** – Pour que la stabilité numérique puisse être améliorée, la recherche en faisceau utilise un objectif normalisé, souvent appelé l'objectif de log-probabilité normalisé, défini par :

$$\text{Objectif} = \frac{1}{T_y^\alpha} \sum_{t=1}^{T_y} \log \left[p(y^{<t>} | x, y^{<1>}, \dots, y^{<t-1>}) \right]$$

Remarque : le paramètre α est souvent comprise entre 0.5 et 1.

□ **Analyse d'erreur** – Lorsque l'on obtient une mauvaise traduction prédite \hat{y} , on peut se demander la raison pour laquelle l'algorithme n'a pas obtenu une bonne traduction y^* en faisant une analyse d'erreur de la manière suivante :

Cas	$P(y^* x) > P(\hat{y} x)$	$P(y^* x) \leq P(\hat{y} x)$
Cause	Recherche en faisceau défectueuse	RNN défectueux
Remèdes	Augmenter la largeur du faisceau	<ul style="list-style-type: none"> - Essayer une différente architecture - Régulariser - Obtenir plus de données