A REINFORCEMENT LEARNING-DRIVEN TRANSLATION MODEL FOR SEARCH-ORIENTED CONVERSATIONAL SYSTEMS

Wafa Aissa, Laure Soulier, Ludovic Denoyer

Sorbonne Université, CNRS, Laboratoire d'Informatique de Paris 6, F-75005 Paris, France

SORBONNE UNIVERSITÉ CRÉATEURS DE FUTURS DEPUIS 1257

MLIA Machine Learning & Deep Learning for Information Access

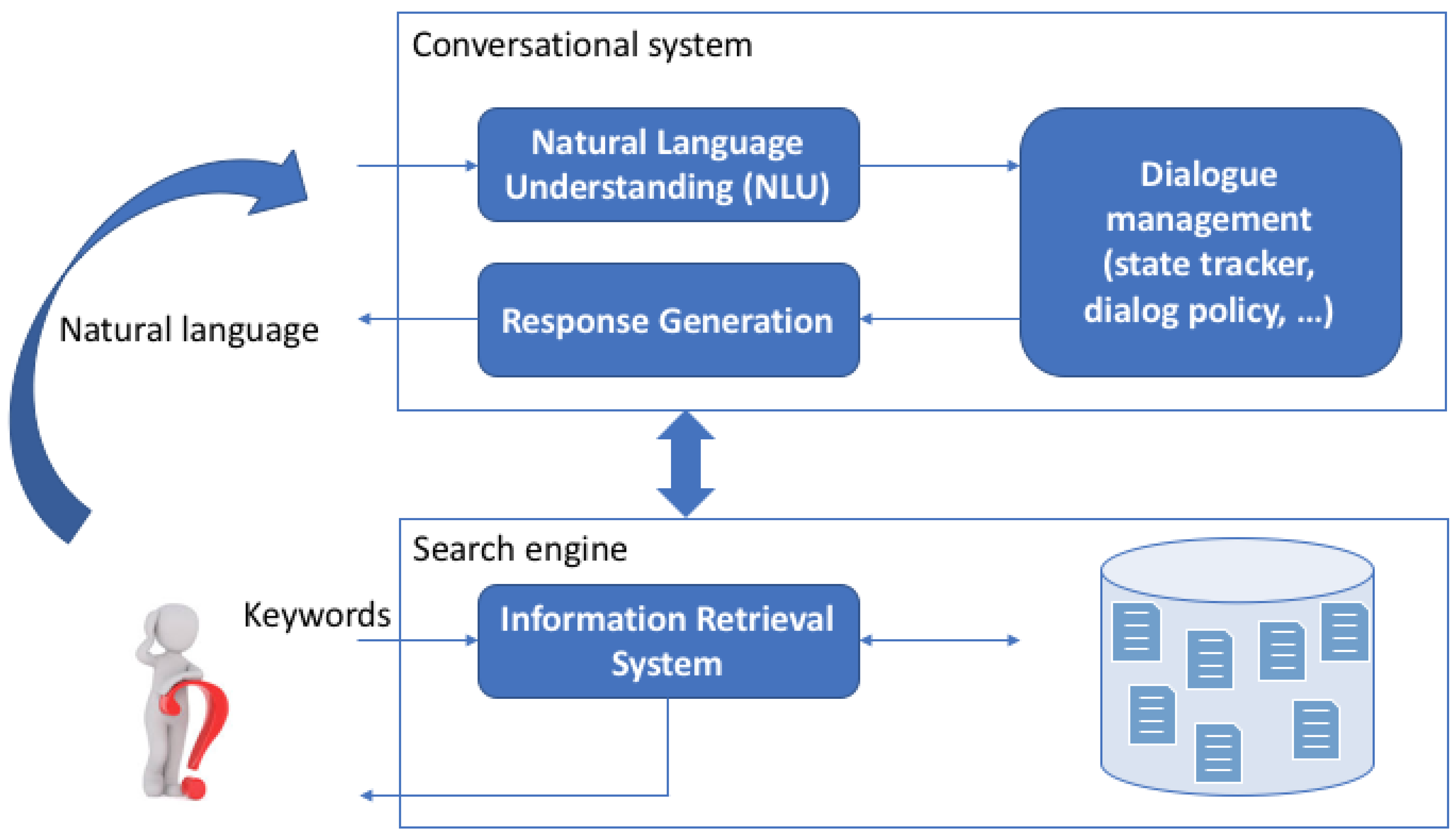
MOTIVATION, BACKGROUND AND RESEARCH QUESTIONS

From conversational systems...

- Chit-chat conversational systems: simple conversations [1]
- Task-oriented conversational systems: closed world (slot-filling patterns, KB extractions, ...) [2,3]
- Users interact in natural language

... to search-oriented conversational systems

- "Open world"
  - Information needs are often ambiguous, vague, exploratory
  - Large document datasets, structured and unstructured information sources
- Information needs expressed through keywords



CHALLENGE

Understand users' information needs expressed in natural language to identify relevant documents

- Build keyword-based queries from natural language expressions
- "End-to-end" approach directly dealing with the NL expression

Related work

- Neural Machine translation models
  - Principle: Encoding the input in a latent representation space and decoding its latent representation in the target language
  - In our context: [4,5]
- Reinforcement learning approaches
  - Principle: Driving the approach by the task
  - In our context: [6]

REINFORCEMENT LEARNING-DRIVEN TRANSLATION MODEL

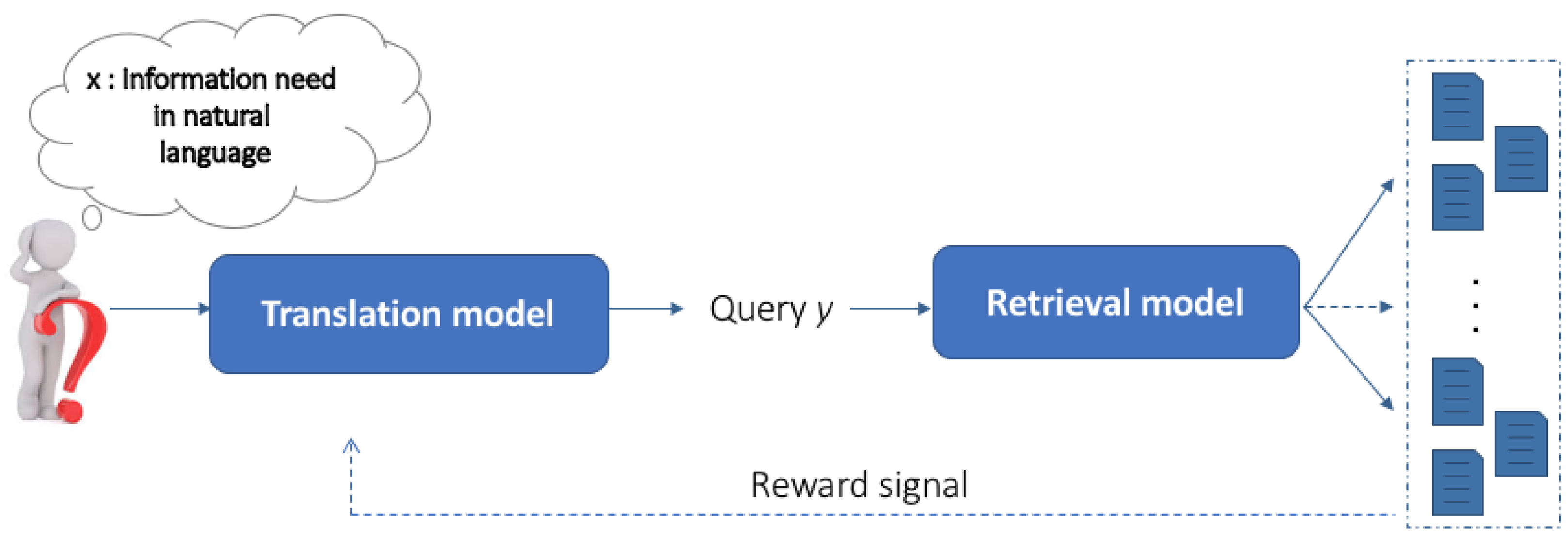
Notations

- $x = x_1, ..., x_i, ..., x_n$: NL user's information need.
- $y \in \{0,1\}^n$: Key-word query.
- $y_i = 1$ if x_i exists in query y and 0 otherwise.

Problem formulation

f_θ: estimate the probability $p(y|x)$ of generating the keywords y given its natural language expression x, conditioned by $y_{<i}$.

$$p(y|x) = \prod_{y_i \in y} p(y_i | y_{<i}, x)$$



PRE-TRAINING WITH SUPERVISED MACHINE TRANSLATION

- NL-query pairs $D = \{(x^1, y^1), ..., (x^k, y^k), ..., (x^N, y^N)\}$
- Likelihood:

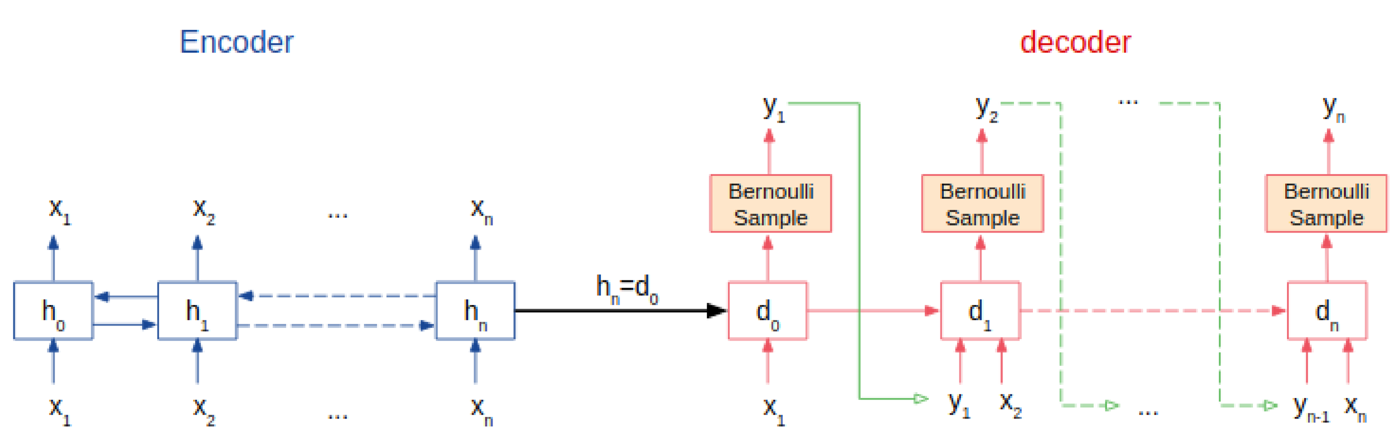
$$L_{SMT} = \sum_{(x^k, y^k) \in D} log(f_\theta(x^k)) \ with \ f_\theta(x^k) = \sum_{y_i^k \in y^k} log(p(\hat{y}_i^k = y_i^k | \hat{y}_{<i}^k, x^k))$$

REFINING WITH REINFORCEMENT LEARNING

- Reward: $R(\hat{y}) = MAP(\hat{y}, \mathcal{D}_x)$ based on documents ranking using y.
- Actions: $\{0, 1\}$ or $\{$select, discard$\}$
- Likelihood:

$$L_{RL}(\theta) = arg \max_{\theta} \mathrm{E}_{\substack{(x;\mathcal{D}_x) \in GT \\ \hat{y} \sim f_\theta(x)}} [R(\hat{y}) - \bar{R}]$$

NEURAL ARCHITECTURE



- Each element x_i of x is modeled through word embeddings w_{x_i}
- Bi-directional LSTM encoder (hidden state h_n)
- LSTM decoder (input: hidden vector h_n, the current word x_i, and a binary indicator y_{i-1} expressing whether previous word x_{i-1} has been selected or not)

EVALUATION

Datasets

- TREC Robust 2004: 250 NL-query pairs, 15.333 documents
- TREC Web 2000, 2001: 100 NL-query pairs, 11.47 documents

Baselines

- NL: the NL information need, TREC descriptions
- Q: original TREC titles
- Qbin: the binary formulated queries
- Random: randomly select 3 words from NL
- SMT: supervised machine translation
- RL: standalone reinforcement learning

Retrieval effectiveness

Baseline	TREC Robust(2004)		TREC Web (2000-2001)	
	MAP	*%Chg*	MAP	*%Chg*
NL	0.08925	+15.25% ***	0.15913	+12.88% *
Q	0.09804	+4.92%	0.16543	+8.58%
Q bin	0.08847	+16.26% *	0.17402	+3.22%
Random	0.01808	+468.91% ***	0.04060	+342.44% ***
SMT	0.06845	+50.27% ***	0.08891	+102.04% ***
RL	0.08983	+14.51% ***	0.16474	+9.04%
SMT+RL	**0.10286**		**0.17963**	

- Low results for **NL**: benefit of using keywords.
- **SMT+RL** overpasses **SMT**: benefit of reinforcement learning.
- **RL** baseline achieves relatively good retrieval performances.
- Reinforcement learning techniques are more effective with pre-training.

Examples of queries

- what are new methods of producing steel?
  - Original query (Q): steel producing
  - Binary query (Qbin): producing steel
  - Our model (SMT+RL):new methods of producing steel
- what are the advantages and or disadvantages of tooth implant?
  - Original query (Q): implant dentistry
  - Binary query (Qbin): implant
  - Our model (SMT+RL):advantages disadvantages tooth implant

CONCLUSION

- A selection model translating users' information need in NL into keyword queries
- Based on supervised machine translation and reinforcement learning.
- Evaluation on two TREC datasets and promising results in terms of effectiveness.

Perspectives: Totally skipping the query formulation step and designing retrieval models dealing with NL expressions.

REFERENCES

[1] Ritter, Alan, Colin Cherry, and William B. Dolan (2011). "Data-driven Response Generation in Social Media"

[2] Bordes, Antoine and Jason Weston (2016). 'Learning End-to-EndGoal-Oriented Dialog'

[3] Dhingra, Bhuwan et al. (2017). "Towards End-to-End Reinforcement Learningof Dialogue Agents for Information Access"

[4] Song, Hyun-Je, A-Yeong Kim, and Seong-Bae Park (2017). "Translation ofNatural Language Query Into Keyword Query Using a RNNEncoder-Decode"

[5] Yin, Zi, Keng-hao Chang, and Ruofei Zhang (2017). "DeepProbe: InformationDirected Sequence Understanding and Chatbot Design via Recurrent NeuralNetworks"

[6] Nogueira, Rodrigo and Kyunghyun Cho (2017). "Task-Oriented QueryReformulation with Reinforcement Learning"