Supplementary Material: Interactively Learning to Summarise by Combining Active Preference Learning and Reinforcement Learning

Yang Gao, Christian M. Meyer, Iryna Gurevych
Ubiquitous Knowledge Processing Lab (UKP-TUDA)
Department of Computer Science, Technische Universität Darmstadt
https://www.ukp.tu-darmstadt.de/

Abstract

In this document, we detail the derivation of the gradient $\nabla_w \mathcal{L}^{SPPI}(w|x)$ for EMDS in Section 1, provide example summary pairs of Unc and SBT in Section 2, and present the performance of APL without prior ranking (i.e. by letting $\alpha = 1$) in Section 3.

1 Adjusting SPPI to EMDS

Here we derive the first order derivative for $\mathcal{L}^{SPPI}(w|x)$. The original SPPI algorithm cannot be directly applied to EMDS, because each output y in EMDS is a summary, and constructing a summary involves making a sequence of actions, namely selecting a sequence of sentences to add to the summary. We extend the gradient descent in original SPPI to the REINFORCE-style policy gradient update (Williams, 1992), similar to the extension performed in (Kreutzer et al., 2017). The difference is that, in (Kreutzer et al., 2017) a neural network is used to approximate the utilities of summaries, but in our model we use the linear approximation as in original SPPI.

First, from Eq. (1) in the main paper, we have:

$$\nabla_w \mathcal{L}^{SPPI}(w|x) = \sum_{y_i, y_j \in \mathcal{Y}(x)} \Delta_x(\langle y_i, y_j \rangle) \; \nabla_w p_w(\langle y_i, y_j \rangle | x) \quad (1)$$

To obtain $\nabla_w p_w(\langle y_i, y_j \rangle | x)$, we rewrite p_w (see Eq. (2) in the main paper) as follows:

$$\begin{aligned} & p_w(\langle y_i, y_j \rangle | x) \\ &= \frac{\exp[w^\intercal(\phi(y_i|x) - \phi(y_j|x))]}{\sum_{y_p, y_q \in \mathcal{Y}(x)} \exp[w^\intercal(\phi(y_p|x) - \phi(y_q|x))]} \\ &= \frac{\exp[w^\intercal \phi(y_i|x)]}{\sum_{y_p \in \mathcal{Y}(x)} \exp[w^\intercal \phi(y_p|x)]} \cdot \frac{\exp[-w^\intercal \phi(y_j|x)]}{\sum_{y_q \in \mathcal{Y}(x)} \exp[-w^\intercal \phi(y_q|x)]} \\ &= p_w(y_i|x) \cdot p_{-w}(y_j|x). \end{aligned} \quad (2)$$

Note that in Eq. (2) we slightly abuse the notation of p_w, so as the build the connection between pairwise selection probability and single summary selection probability: we can see that the probability of selecting pair $p_w(\langle y_i, y_j \rangle | x)$ is the product of two single summary selection probabilities: $p_w(y_i|x)$, the probability of selecting y_i in cluster x using weights vector w, and $p_{-w}(y_j|x)$, the probability of selecting y_j according to weights $-w$. Based on this split, we have

$$\begin{aligned} & \nabla_w p_w(\langle y_i, y_j \rangle | x) \\ &= p_{-w}(y_j|x) \nabla_w p_w(y_i|x) + p_w(y_i|x) \nabla_w p_{-w}(y_j|x). \end{aligned} \quad (3)$$

In the reminder of this section, we assume a unique cluster x is given, and thus we omit x in our equations to ease the presentation. Now we extend $\nabla_w p_w(y_i|x)$ in detail; $\nabla_w p_{-w}(y_j|x)$ can be obtained similarly. We assume y_i is a summary consisting of K sentences $(y_i^1, y_i^2, \cdots, y_i^K)$. Then we have

$$\begin{aligned} \nabla_w p_w(y_i) &= p_w(y_i) \nabla_w \log p_w(y_i) \\ &= p_w(y_i) \sum_{k=1,\cdots,K} \nabla_w \log p_w(y_i^k), \end{aligned} \quad (4)$$

where $p_w(y_i^k)$ is the probability of concatenating sentence y_i^k to the existing draft summary $(y_i^1, \cdots, y_i^{k-1})$. $p_w(y_i^k)$ is also a Gibbs sampling strategy:

$$p_w(y_i^k) = \frac{\exp[w^\intercal \phi(yk_i)]}{\sum_q \exp[w^\intercal \phi(yq_i)]} \quad (5)$$

where yk_i is the resulting summary of adding sentence y_i^k to the existing summary $(y_i^1, \cdots, y_i^{k-1})$, namely $(y_i^1, \cdots, y_i^k)$. Similarly, yq_i is the resulting summary of adding a sentence q into the existing summary $(y_i^1, \cdots, y_i^{k-1})$, namely

$(y_i^1, \cdots, y_i^{k-1}, q)$, where q ranges over all available sentences that have not been included in the current summary yet. $\phi(y)$ is the vector representation of a draft summary y. Let $a_k = \exp[w^\intercal \phi(yk_i)]$ and $c = \sum_q \exp[w^\intercal \phi(yq_i]$, we can thus rewrite Eq. (5) as $p_w(y_i^k) = a_k/c$. Then we can derive the derivative of $\log p_w(y_i^k)$ as follows:

$$\begin{aligned}\nabla_w \log p_w(y_i^k) &= \frac{c}{a_k} \cdot \frac{c \nabla_w a_k - a_k \nabla_w c}{c^2} \\ &= \frac{\phi(yk_i)c - \sum_q a_q \phi(yq_i)}{c} \\ &= \phi(yk_i) - \sum_q p_w(q)\phi(yq_i).\end{aligned} \tag{6}$$

Computation in Eq. (6) is expensive because it needs to compute the probability $p_w(q)$ for every unused sentence q. Because our 100-word summaries usually only contain 3-5 sentences, q is almost as big as the number of all sentences in a document cluster (hundreds to even a few thousand). In addition, Eq. (6) has to be computed for K times for computing $\nabla_w p_w(y_i)$ (see Eq. (4)), and another K times for computing $\nabla_w p_{-w}(y_j)$. This explains the high computational time of SPPI for EMDS. By combining Equations (1), (3), (4) and (6) altogether, the gradient $\nabla_w \mathcal{L}^{SPPI}$ can be obtained.

2 Example Summary Pairs

Below we present some example summary pairs presented by Unc and SBT. All summaries meet the 100-word limit, from a randomly selected topic (d068f in DUC'02). The topic is about the removal of Checkpoint Charlie, the Berlin Wall border post that symbolised the Cold War.

SBT Example Pair 1:

- The famed Allied checkpoint by the Berlin Wall was closed with an elaborate ceremony that brought together the top diplomats from the Germanys and the four World War II Allies. Patrick Gainey took pictures for the U.S. Army. Baker, Soviet Foreign Minister Eduard Shevardnadze and the foreign ministers from France, Britain and the two Germanys each heralded the end of the checkpoint as a symbol of change. The Soviet Union said today that a united Germany can join NATO after a five-year transition period during which all Soviet and U.S. troops would leave the country.
- "He drove a half-mile into the west side just to make sure. West Germany now is a mainstay of NATO and East Germany is a member of the Soviet-led Warsaw Pact. Checkpoint Charlie went up in 1961 in the middle of the Friedrichstrasse boulevard after Communist East Germany erected the Berlin Wall to choke off a flood of refugees to the enclave of West Berlin. "This guy had driven a cement truck with his girlfriend and baby through the checkpoint," said Gainey, a U.S. Army photographer. Ed McCarthy said he believes it is destined for a museum. Bratke just ran.

SBT Example Pair 2:

- Checkpoint Charlie, the Berlin Wall border post that symbolized the Cold War, was hoisted into history today. The United States wants a united Germany to be a member of the North Atlantic Treaty Organization and the Soviet Union has been staunchly opposed. It was part of my home." West Germany now is a mainstay of NATO and East Germany is a member of the Soviet-led Warsaw Pact. The famed Allied checkpoint by the Berlin Wall was closed with an elaborate ceremony that brought together the top diplomats from the Germanys and the four World War II Allies. Bratke just ran.
- "I ran as fast as I could," he said from his home in West Berlin. Genscher thanked Moscow for reform policies that had ended "a time of trouble" in Berlin and opened the door to unity. There are about 380,000 Soviet troops in East Germany, while a quarter-million U.S. troops are stationed in West Germany along with tens of thousands of British and French soldiers. Patrick Gainey took pictures for the U.S. Army. It was part of my home." He was shot in the neck." "I saw Kennedy, Reagan and Carter give speeches here," she said. Bratke just ran.

Unc Example Pair 1:

- Large sections of the Berlin Wall are being dismantled daily, and U.S. officials decided to time the removal of Checkpoint Charlie with the foreign ministers' meeting. "For 29 years Checkpoint Charlie embodied the Cold War," U.S. Secretary of State James A. Baker

III told a crowd of invited guests and journalists on Berlin's Friedrich Street. Checkpoint Charlie, the famed Allied border crossing on the west side of the Berlin Wall, was lifted into the sky by a giant crane Friday, placed gently onto a flatbed truck and consigned to history. It was part of my home."

- Checkpoint Charlie, the famed Allied border crossing by the Berlin Wall, was to be hauled away Friday. Maik Polster was a stern-faced member of the East German secret police. The proposal was outlined by Soviet Foreign Minister Eduard Shevardnadze during international talks in East Berlin on the strategic future of a united Germany. This is wonderful for history, but it is sad for me. Shevardnadze, the first Soviet foreign minister to visit West Berlin, noted that the checkpoint was vanishing on the 49th anniversary of the Nazi invasion of the Soviet Union. He was taking his certification examinations this week.

Unc Example Pair 2:

- The public was not invited to the ceremony on Friedrich Street, the avenue blocked by Checkpoint Charlie and its once-feared East German counterpart nearby. "We always hoped that one day Checkpoint Charlie would no longer be needed, and now that day has arrived. Checkpoint Charlie stood just 20 yards from an East German checkpoint and was the only place between the Berlins that could be crossed on foot by non-Germans. Soviet Foreign Minister Eduard A. Shevardnadze attended the ceremony in a gesture of reconciliation with the Western members of the wartime alliance : Britain, France and the United States.

- Soldiers from three nations locked the doors of Checkpoint Charlie today and a crane whisked away the Cold War relic that symbolized a divided world. The ceremony was closed to the public but not to the residents of the buildings that line Friedrich Street, which had been divided by the Berlin Wall since 1961. Bratke is now a West Berlin engineering student. Shevardnadze, the first Soviet foreign minister to visit West Berlin, noted that the checkpoint was vanishing on the 49th anniversary of the Nazi invasion of the Soviet Union. U.S. Army spokesman Sgt. "I still can not believe it."

	RND		SBT		Unc		J&N	
Oracle	τ	ρ	τ	ρ	τ	ρ	τ	ρ
Query budget $T = 10$, $\alpha = 1$:								
PO	.074	.110	.168	.249	**.203**	**.300**	.081	.119
CNO-0.1	.007	.105	.159	.235	**.179**	**.265**	.081	.121
CNO-0.3	.058	.086	.092	.136	**.146**	**.218**	.065	.097
LNO-0.3	.044	.065	.117	.173	**.136**	**.202**	.063	.094
LNO-1	.027	.039	.070	.104	**.058**	**.086**	.028	.041
Query budget $T = 100$, $\alpha = 1$:								
PO	.192	.283	.329*	.476*	**.357***	**.511***	.185	.273
CNO-0.1	.171	.253	.267*	.391*	**.329***	**.474***	.170	.252
CNO-0.3	.123	.193	.154	.226	**.259***	**.379***	.118	.175
LNO-0.3	.121	.180	.211	.312	**.255***	**.373***	.133	.199
LNO-1	.048	.072	.083	.124	**.172**	**.254**	.056	.084
Baseline: $\tau = .206$, $\rho = .304$								

Table 1: Performance of multiple APL algorithms (columns) using different oracles and query budgets (rows). The baseline is the purely prior ranking ($\alpha = 0$), without any interaction ($T = 0$). All results except the baseline are averaged over 50 document clusters in DUC'04. Asterisk: significant advantage over the baseline.

3 Performance of APL without Prior Knowledge

Table 1 presents the ranking results of different APL strategies, without using prior knowledge HU (i.e. $\alpha = 1$). When the query budget T is 10, no APL is able to surpass the baseline. When the budget is 100, only Unc and SBT are able to significantly improve the baseline when interacting with low-noise oracles PO and CNO-0.1. These results indicate the importance of trading off between prior and posterior information to produce rankings, especially when the budget is low and/or the oracle's noise level is high.

References

Julia Kreutzer, Artem Sokolov, and Stefan Riezler. 2017. Bandit Structured Prediction for Neural Sequence-to-Sequence Learning. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (ACL 2017), Volume 1: Long Papers, July 30 - August 4, Vancouver, Canada*, pages 1503–1513.

Ronald J. Williams. 1992. Simple statistical gradient-following algorithms for connectionist reinforcement learning. *Machine Learning*, 8:229–256.