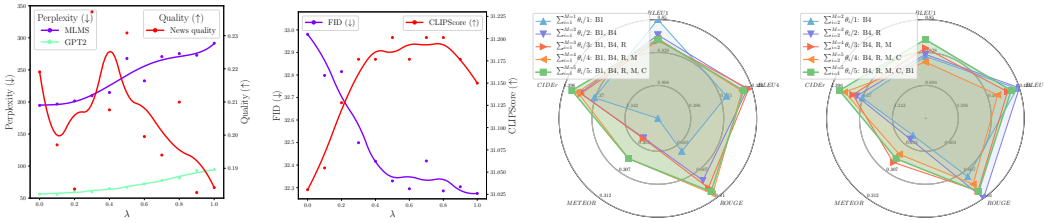
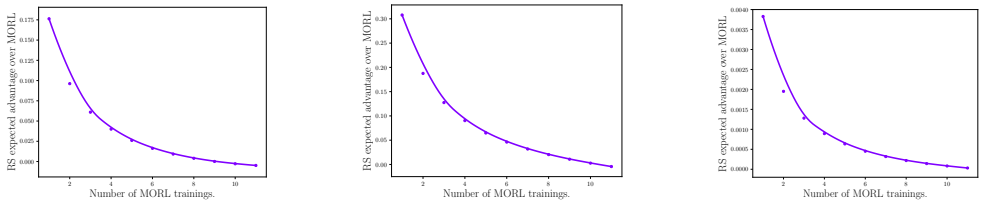
Table 1: Examples of generated summaries about vampire vigilantes in Malawi news, for $(1-\lambda)\cdot\theta_1+\lambda\cdot\theta_2$ interpolating between θ_1 and θ_2 fine-tuned on R_1 (evaluating completeness) and R_2 (evaluating faithfulness).

λ	Generation
0.0	Malawi vigilante mobs have been killing people suspected of being vampires, prompting the UN and US embassy to declare no-go zones in the affected'
0.1	Malawi vigilante mobs have been killing people suspected of being vampires, prompting the UN and US embassy to declare no-go zones in the affected
0.2	Malawi vigilante mobs have killed nine people since mid-September, prompting the United Nations and the US embassy to declare some parts of the country no-
0.3	Malawi vigilante mobs have killed nine people, prompting the UN and US embassy to declare parts of the country no-go zones due to widespread
0.4	Malawi vigilante mobs have killed nine people, prompting the UN and US embassy to declare some parts of the country no-go zones.
0.5	Malawi vigilante mobs have arrested and killed suspected vampires, prompting the UN and US embassy to declare no-go zones and President Peter Muth
0.6	Malawi vigilante mobs have arrested and killed suspected vampires, prompting the UN and US embassy to declare no-go zones.
0.7	Malawi vigilante mobs have arrested suspected vampires, resulting in deaths and prompting the UN and US embassy to declare no-go zones.
0.8	Malawi vigilante mobs have arrested suspected vampires, resulting in deaths and violence.
0.9	Malawi vigilante violence has caused widespread panic and death, prompting authorities to arrest suspected members and investigate the belief in vampirism.
1.0	Malawi vigilante violence has caused widespread panic and death.



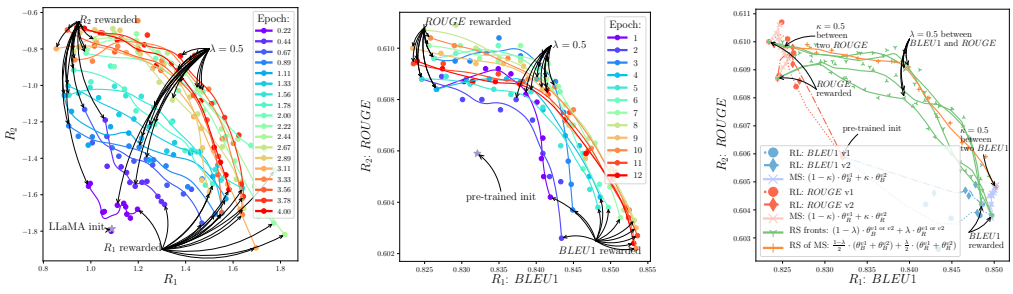
(a) News summary. (b) Text-to-image. (c) Captioning: B1 first. (d) Captioning: B4 first.

Figure 1: (a) evaluates summaries generated by λ-interpolated LLMs in terms of perplexity (by MLMS or GPT2) or quality by this news quality model. (b) evaluates images generated by λ-interpolated diffusion models in terms of realism by FID or text alignment by CLIPScore. The spider maps (c,d) uniformly average $1 \leq M \leq 5$ weights for captioning, where θ_1 is fine-tuned on BLEU1 (B1), θ_2 on BLEU4 (B4), θ_3 on ROUGE (R), θ_4 on METEOR (M) and θ_5 on CIDEr (C). To show different combinations among the $\binom{5}{M}$ possible, we iterate in a clockwise direction starting in (c) from $i=1$ (always including θ_1) and in (d) from $i=2$ (always including θ_2).



(a) News summary. (b) Reddit summary. (c) Captioning.

Figure 2: Expected reward advantage of RS (always requiring only 2 trainings) over MORL (with M trainings), defined as $\mathbb{E}_{\hat{\mu}\sim Unif(0,1)}\left[max_{\lambda\in\Lambda}\hat{R}_{\hat{\mu}}(\theta^{RS}_{\lambda}) - \mathbb{E}_{\Lambda_M}\left[max_{\mu\in\Lambda_M}\hat{R}_{\hat{\mu}}(\theta^{MORL}_{\mu})\right]\right]$, where $\hat{R}_{\hat{\mu}} = (1-\hat{\mu})\times R_1 + \hat{\mu}\times R_2$ is the user reward for user linear preference $\hat{\mu}$ sampled uniformly between 0 and 1, $\Lambda = \{0, 0.1, ..., 1.0\}$ is the set of the 11 possible values for λ, and where the expectation for the MORL term is over the $\binom{11}{M}$ possible combinations Λ_M of M elements from Λ (representing the M linear weightings μ used for MORL training). We observe that MORL matches RS only for M sufficiently big.



(a) News summary. (b) Captioning. (c) RS of MS.

Figure 3: (a,b) show how RS's fronts evolve over the course of fine-tuning, and confirms the LMC even when doubling the number of training epochs (previously 2 for summary and 6 for captioning). (c) enriches previous Figure 10.b showing that RS (rewarded soups) and MS (model soups) are complementary. We consider two fine-tunings for BLEU1 (v1 and v2), that we κ-interpolate in MS. The orange line λ-interpolates the MS for BLEU1 and the MS for ROUGE with $\kappa = 0.5$. Overall, RS reveals a large front while MS mostly reduces variance.