Interactive Open-Domain Story Generation

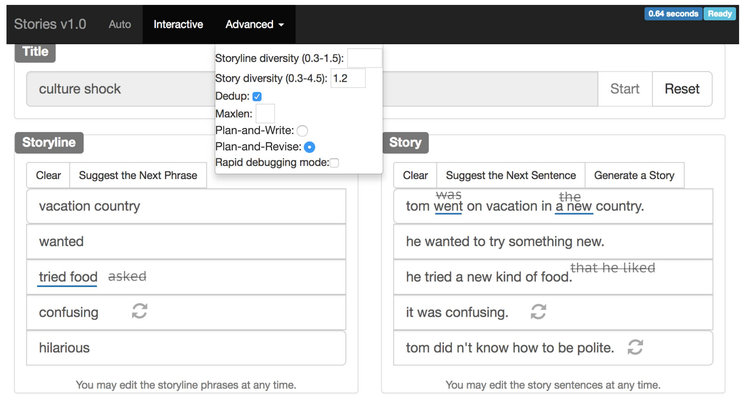
Seraphina Tarrant, Haining Feng, Nanyun Peng

Motivation

- *Can human-machine collaboration improve open-domain neural story generation?*
- *Can it improve specific story aspects, as well as overall quality?*

Previous approaches to human-machine collaboration offer limited interaction; they focus on turn-taking and do not enable iteration. We design a system that enables human interaction at multiple stages of the process: story-planning, story-writing diversity controls*, and model-selection.

Sample Interaction

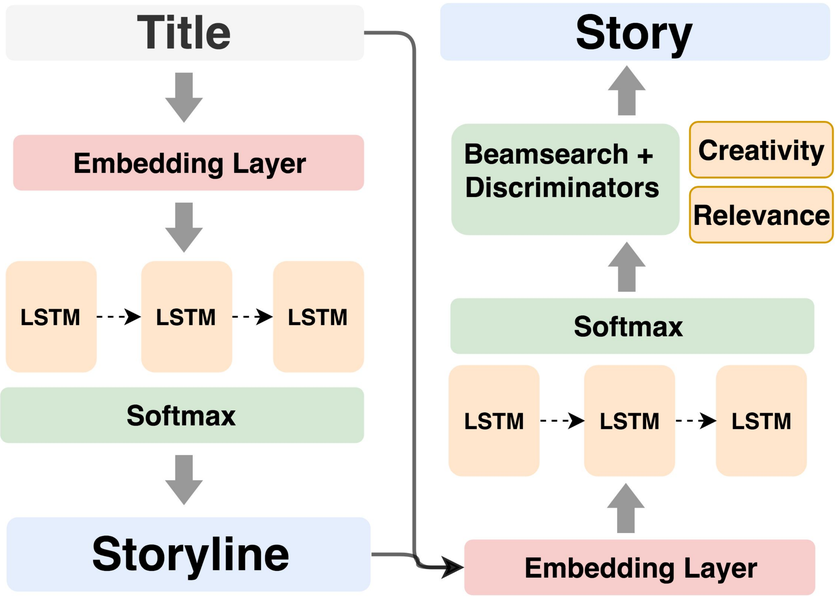


Demo UI annotated with sample user actions

We conduct user studies for baseline turn-taking systems & explore multiple interaction scenarios: ***all-interaction***, ***story-only***, ***storyline-only***, and ***diversity-only*** variations under 5-minute time constraints. We task some users with improving specifically one of Relevance, Creativity, or Causal-Temporal Coherence.

**diversity controls* are softmax temperatures, which control the unusualness of system generations

System



System Details

We adapt the Plan-and-Write system; a storyline-planning to story-generation pipeline trained on ROCStories, to enable interaction at the story-planning stage. We additionally trial a Title-to-Story baseline (no planning) & create a new Plan-and-Revise system by incorporating discriminators for Relevance & Creativity.

Plan-and-Revise modifies the language model decoding objective as below, where s_k is the scoring function learnt by the discriminator and λ_k is a learned weight coefficient:

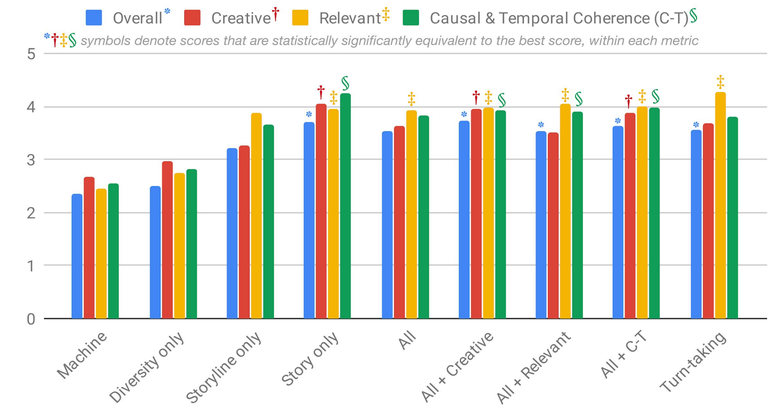
$$f_\lambda(\mathbf{x}, \mathbf{y}) = \log(P_{\text{lm}}(\mathbf{y}|\mathbf{x})) + \sum_k \lambda_k s_k(\mathbf{x}, \mathbf{y})$$

Thus the system leverages both *automated* and *human* collaborative discriminators, and can isolate the contribution of each.

Metrics

Self-reported Subjects self-report on: engagement, satisfaction with story, & perception of story quality.
Independent Ranking Independent human judges are asked to rank all stories from 1-5 under eight experiment conditions for Overall Quality, Relevance, Creativity, and Causal-Temporal Coherence.

Results



- humans tasked with improving a *specific* story aspect **are** successful at doing so.
- interaction at both *planning* and *writing* stages improves story quality 10-50% over less interactive baselines.
- additional interaction increases user self-reported satisfaction in all areas (refer to tables in paper).

Code and data available at:
https://github.com/seraphinatarrant/plan-write-revise

Blog Post: https://tinyurl.com/pwr-naacl-2019
Paper: *Plan, Write, and Revise: an Interactive System for Open-Domain Story Generation*, on ArXiv.