Dynamically Revising Neural Networks via Continual Knowledge Feedback

Dr. Mohit Bansal, UNC Chapel Hill

Driving Applications: Lifelong Learning Machines (L2M)

L2M Motivation and Vision

- L2M Core Capabilities Addressed:
  - **Continual Learning**: Model learns its neural architecture continually as new tasks keep coming in, without sacrificing previously-learned knowledge (and with **interpretability for safety**).
  - **Adaption to New Circumstances**: Generalization to new, unseen domains/ tasks via external, general knowledge reinforcement rewards; and via neural memory cell self-learning.
  - **Selective/Balanced Plasticity**: Self-selection multi-task models that decide which sub-model parts/layers to share/update across tasks vs. retain as fixed/ private; and via orthogonality-constrained continual architecture learning that retains previously-learned knowledge.
- Overall Outcome: Feedback-based, dynamically-adapting (self-correcting), commonsense-capable, lifelong learning models that iteratively+selectively learn from their errors and from external general knowledge, and also generalize well to new unseen test-time scenarios, without sacrificing previously-learned skills.

Innovation w.r.t. Previous Work

- Most current ML models employ static parameter learning and a manually-defined architecture.
- In contrast, our model can update its whole neural model architecture (based on human feedback/corrections and external commonsense knowledge), allowing for much stronger adaptability, w/o hurting previous task accuracy.
- We also allow free-form verbal feedback and predict user satisfaction (via sentiment and prosody emotion analysis).
- We add a dynamic common sense graph extracted from external knowledge bases via reinforcement learning (RL) and relevance to current conversation.
- New knowledge gets continually added, and gets assimilated with previous parameters as a lifelong-growing common sense, similar to how humans learn.

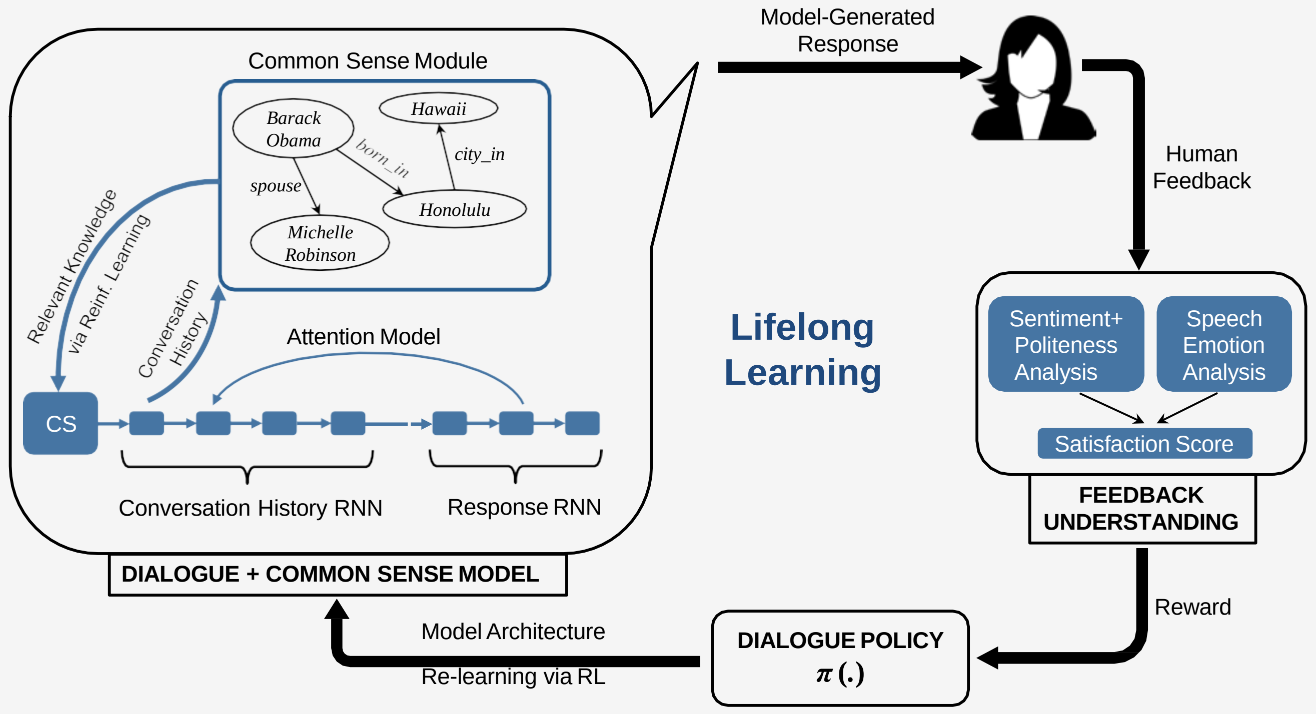
Impact and Applications

- Our dynamically-revised neural network will continually learn from its errors (w/o forgetting previous errors) and also have human-like common sense.
- This new class of lifelong learning models will bring substantial improvements to several important AI tasks, e.g., robotic task instruction execution and collaboration (for troubleshooting, navigation, assembly, etc.)
- Impactful in DoD-relevant remote scenarios (e.g., building on fire or conflict site), where robot harnesses corrective feedback as well as external common sense to regularly update itself.

Technical Approach

- **Research Challenge 1: Neural Architecture Revisions (Dialogue Models)**
- Policy-RL and DAG-sampling controllers to update model architecture, based on tree recurrent neural networks (RNNs) and memory cell, #layers, #nodes

- **Research Challenge 2: Free-form User Feedback**
- RL rewards on sentiment analysis, politeness [Aubakirova & Bansal, 2016], speech emotion (pitch, rate, pauses, energy, tilt). Multimodal facial expression.



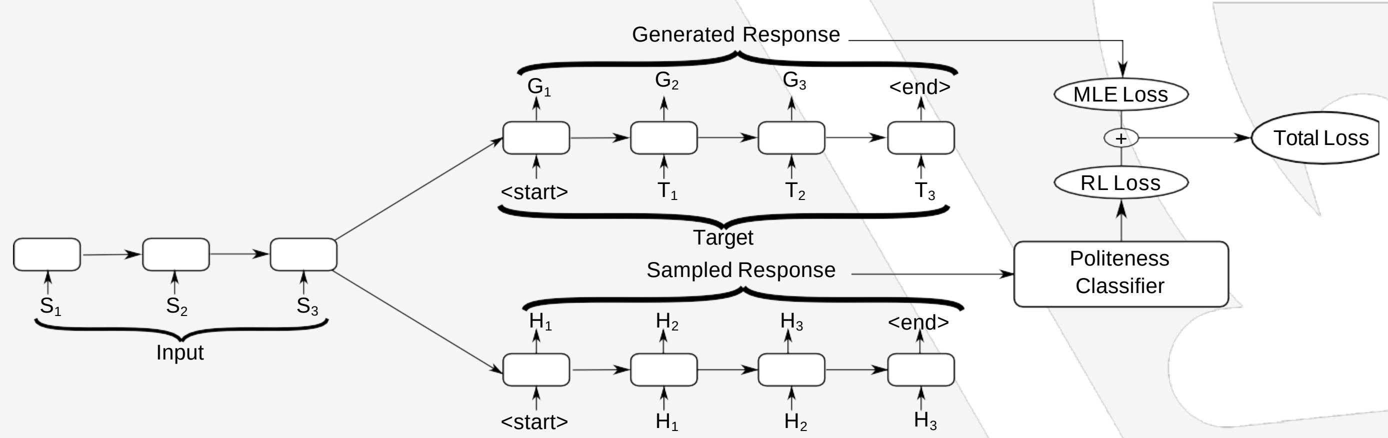
- **Research Challenge 3: Commonsense:**
- Commonsense relevant to dialogue is extracted from multiple knowledge graphs, via subgraph embedding matching, multi-source RL reconciliation.

- **Research Challenge 4: Remembering + Distilling Previous Knowledge**
- RL policy iteratively updates to find optimal long-term policy that maximizes total sum of rewards. Life-long commonsense via continually-growing subgraphs and distillation.

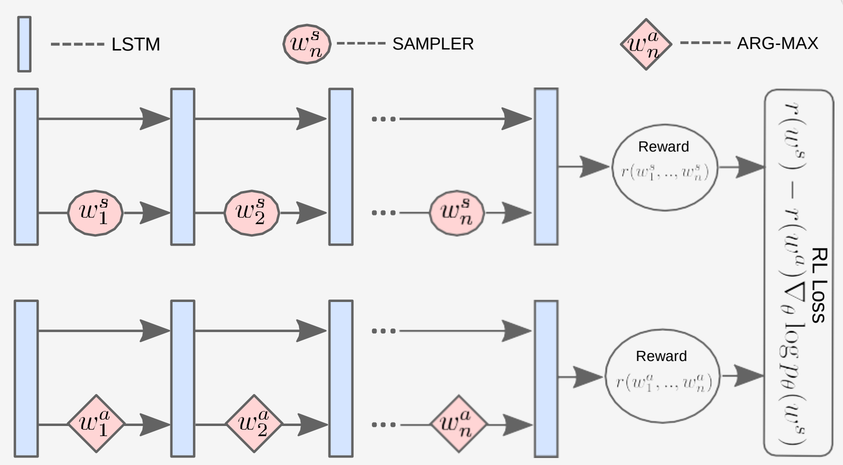
- **Evaluation:** Robotic-instruction navigation/configuration tasks. Metrics: automatic (error rate, retrieval, completion rate) + human (relevance, fluency).

Result 1: Knowledge-Feedback Decoders

- Feedback RL rewards based on external knowledge classifiers encourage model to continually generate more stylistic/logical/salient responses, while maintaining context/history relevance [Niu & Bansal, 2018].

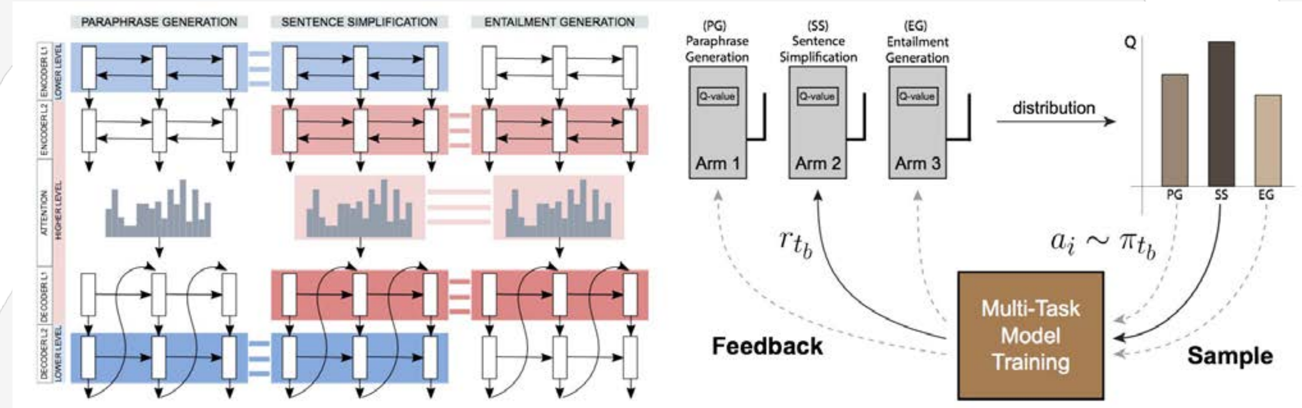


- Multi-reward optimizer, based on multi-task learning, adds general knowledge to improve *test-only generalization* results [Pasunuru & Bansal, 2018].



Result 2: Self-Selection Multi-Task Learning

- We propose a multi-armed bandit approach that dynamically self-learns a schedule of switching between tasks for optimization during multi-task learning (based on rewards from task performance, which can be transfer-learning out-of-domain), instead of the traditional manually-tuned static (fixed) mixing ratio.

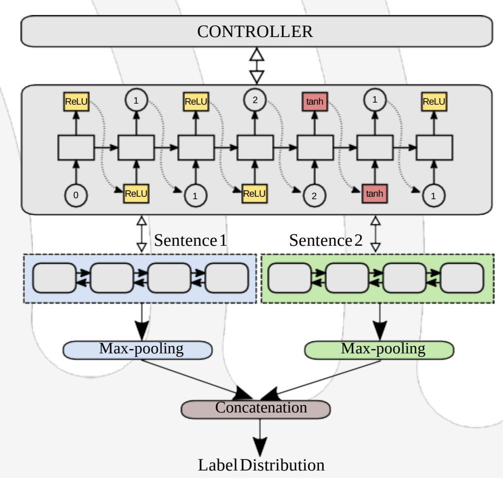


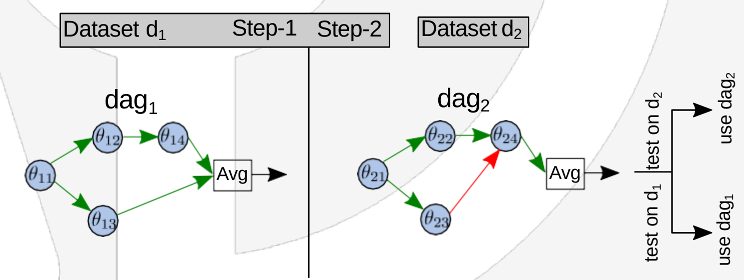
- It can also self-decide which layers of each task to *selectively share* (e.g., higher-level semantic vs. lower-level syntactic) vs. keep unshared/private (to retain task-specific knowledge via selective-plasticity), and also which tasks to share itself for generalization [Guo, Pasunuru, Bansal, 2018].

Models	BLEU	FKGL	SARI
Previous Work			
PBMT-R	18.19	7.59	15.77
Hybrid	14.46	4.01	30.00
EncDecA	21.70	5.11	24.12
DRESS	23.21	4.13	27.37
DRESS-LS	24.30	4.21	26.63
Our Models			
Baseline ⌧	23.72	3.25	28.31
⌧+ Ent.	16.82	2.21	31.55
⌧+ Paraphr.	16.29	2.03	31.71
⌧+Ent+Par	11.86	1.38	32.98
Static Mixing Ratio	11.86	1.38	32.98
Dynamic Mixing Ratio	11.14	1.32	33.22

Result 3: Continual Architecture Search

- Novel '*continual architecture search*' approach (CAS), to dynamically evolve model structure (e.g., RNN memory cell as directed acyclic graph) during sequential training of tasks, while maintaining performance on previously learned tasks (2-step graph-initialization w/ block sparsity and orthogonality conditions), enabling life-long learning [Pasunuru & Bansal, 2018b].





Models	ABSA-L	ABSA-R
Previous Work		
Augenstein (2018)	76.74	67.47
AS Results		
Baseline	77.01	81.20
AS	**78.93**	**83.15**
CAS Results		
CAS step-1	78.86	N/A
CAS step-2	78.24	83.21

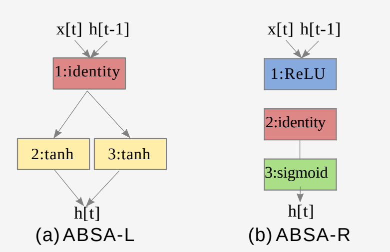


Fig. 1: Learned cell structures during AS.

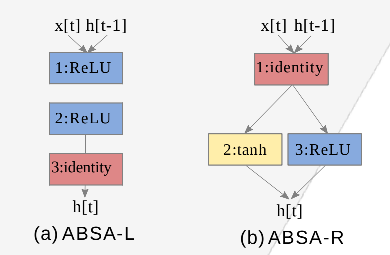


Fig. 2: Learned cell structures during CAS.

References:

1 M. Aubakirova and M. Bansal. Interpreting Neural Networks to Improve Politeness Comprehension. *EMNLP 2016.*
2 T. Niu and M. Bansal. Polite Dialogue Generation Without Parallel Data. *TACL 2018.*
3 R. Pasunuru and M. Bansal. Multi-Reward Reinforced Summarization with Saliency and Entailment. *NAACL 2018.*
4 H. Guo, R. Pasunuru, and M. Bansal. Dynamic Multi-Level Multi-Task Learning for Sentence Simplification. *COLING 2018.*
5 R. Pasunuru and M. Bansal. Continual Architecture Search for Text Classification. *In Submission, 2018.*

THE ELECTRONICS RESURGENCE INITIATIVE

This research was developed with funding from the Defense Advanced Research Projects Agency (DARPA)
The views, opinions and/or findings expressed are those of the author and should not be interpreted as representing the official views or policies of the Department of Defense or the U.S. Government