A Deep-Learning Based Framewrok for T&E as a Continuum

$\Big($A White Paper$\Big)$

Avi Kak

Purdue University

West Lafayette, IN 47907

kak@purdue.edu

Paul Campo

Darkflow

13873 Park Center Road, Fifth Floor, Herndon, VA 20171

pcampo@ewa.com

T&E (Test and Evaluation) is a critical part of practically every system development effort for enterprise-level projects. Before the end-product produced by the effort can be accepted by its intended client, it must pass the T&E criteria. Of all of the different stages of R&D, the one that has so far remained most resistant to automation is T&E. The main problem is that the test cases must still be manually specified and, in highly complex systems, it is more often the case than not that the humans are unable to anticipate all possible test cases relevant to all possible application scenarios for the end-product.

We believe that that the latest advances in Artificial Intelligence (AI) make it an imperative (especially for the R&D related to our national defense projects) that there be a reexamination of the fundamental tenets of the current T&E practices. It is the realization of this imperative that underlies the proposed research that will be carried out by a collaboration between Darkflow and Purdue RVL.

The starting point for our proposed research is the "T&E as a Continuum" paradigm that was first enunciated by Collins and Senechal in the March 2023 issue of The Journal of Test & Evaluation. As the authors argue, the current practice consists of three discrete stages: (1) Contractor Testing (CT), (2) Developmental Testing (DT), and (3) Operational Testing (OT). This serial process results in "late discovery of the issues affecting performance" through OT. And that the issues thus discovered are "extremely challenging to address" since "program planning and funding typically do not incorporate time and resources needed to comprehensively address" such problems. The paradigm "**T&E as a Continuum**" is meant to redress these shortcomings of the serial approach

with a continuous-time collaborative framework that involves both the systems and the mission oriented aspects of R&D along with T&E.

The immediate goal of our research is to map the "T&E as a Continuum" paradigm to a set of collaborating neural networks, with the three main networks representing CT, DT, and OT. The other main network in the system would be for Requirements Verification (RV). The logic in each network would be driven by a graph-theoretic representation of the Entity-Relationship (ER) model for the domain that is relevant to that network. The Entities would represent the main elements of the Requirements Specification for the CT and the DT models. On the other hand, for the OT model, the Entities would represent the metrics that measure mission effectiveness.

The heart of the proposed system would be a supervisory network that tries to detect on a real-time basis inconsistencies between the ER models maintained by the CT, DT, OT, and RV networks. The change detection itself can be done by using a Graph Neural Network (GNN). These have become increasingly important in the deep-learning community for solving tough problems that involve relational data. More specifically, the change detection framework would consist of GNNs embedded in what's known as a Siamese Architecture in which the two relational structures, one a possibly changed version of the other, are supplied as inputs to the two inputs of the Siamese network that is trained to detect changes between the two inputs.

That leaves unaddressed only the problem of how to construct the Entity-Relationship (ER) representations for the CT, DT,, OT, and RV models. Ideally, these would be automatically generated from the Requirements Specification. In our own research, though, we will start with human-specified ER models in order to create and test the rest of the overall architecture. Subsequent effort would address auto-generation of the CT, DT, OT, and RV models from the Requirements Specification documents.

The great beauty of the approach presented above is that it can work with any degree of (initially human-specified) detail in the CT, DT, OT and RV models. As a case in point, let's say that the design engineers specified some N parameters (these could be any subset of all the parameters relevant to the design) to be important for the overall design of the product and for its performance evaluation. The ER models for the CT, DT, OT and RV would only need to know about those N parameters for our system to work. That is, with respect to those N parameters, our system would guarantee consistency between the CT, DT, OT, and the RV models on a continuous-time basis. That itself could serve as a powerful constraint on ensuring the end-to-end coordination between all stages of product development, including T&E, so that we do not have to "settle" for the presented state of the end-product because the program management did make allowances for late-state changes to the design.

End Notes:

- Upon request, the authors will gladly supply the knowledge base used for the ideas presented in this white paper.

- Also available upon request is a record of Purdue RVL's past accomplishments in Deep Learning based architectures.