Functional Object-Oriented Network for Manipulation Learning

David Paulius, Yongqiang Huang, Ahmad Babaeian, Roger Milton, William D. Buchanan, Jeanine Sam and Yu Sun, Ph.D.

University of South Florida, FL, US.

Abstract:

Our work presents a novel structured knowledge representation called the functional object-oriented network (FOON) to model the connectivity of functional-related objects and their motions in manipulation tasks. The graphical model FOON is learned by observing object state change and human manipulations with the objects. Using a well-trained FOON, robots can decipher a task goal, seek the correct objects at the desired states on which to operate, and generate a sequence of proper manipulation motions. FOON comprises of singular units which we coin the term *functional unit* (shown in Figure 1).

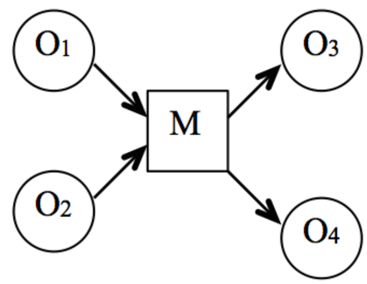

*Figure 1: The basic **functional unit** with two input objects, an interactive motion node, and two output objects.*

We describe FOON's structure, how we capture and represent knowledge, and an approach taken to form a universal FOON with extracted knowledge from online instructional videos. A graph retrieval approach is presented to generate manipulation motion sequences from the FOON to achieve a desired goal, demonstrating the flexibility of FOON in creating a novel and adaptive means of solving a problem using knowledge gathered from multiple sources. Our current focus lies in gathering knowledge in the service domain of cooking. An example of how the knowledge of one cooking video is shown in Figure 2. The results are demonstrated in a simulated environment to illustrate the motion sequences generated from the FOON to carry out the desired tasks. This simulation environment developed in Unity shows how objects will interact with one another to produce a specific outcome using available ingredients and utensils. These motion sequences are carried out by following a task sequence which may not necessarily be knowledge from a single source. Our current network is visualized as Figure 3.

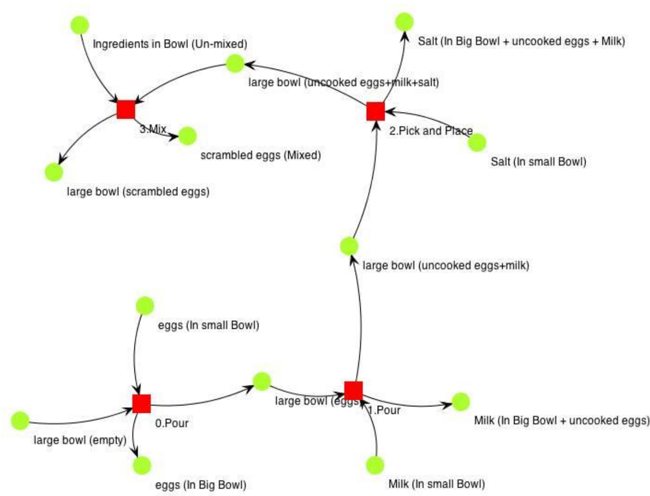

Figure 2: Example of a subgraph demonstrating the procedure to preparing scrambled eggs. Objects are denoted by green circular nodes, while motions are denoted by red square nodes.

We also incorporate the use of network concepts and analyses, namely centrality measures, to determine the important nodes and to find associations between certain objects. These metrics along with motion analyses allow us to prepare for the common case and ensure that object manipulation is perfected especially with those items. We also demonstrate the use of FOON for clustering objects into groups based on their functional-relation, that is, objects which are found to be manipulated in a similar manner. With this information, we hope to make FOON more intelligent such that we can infer the manipulation of objects unknown to FOON.

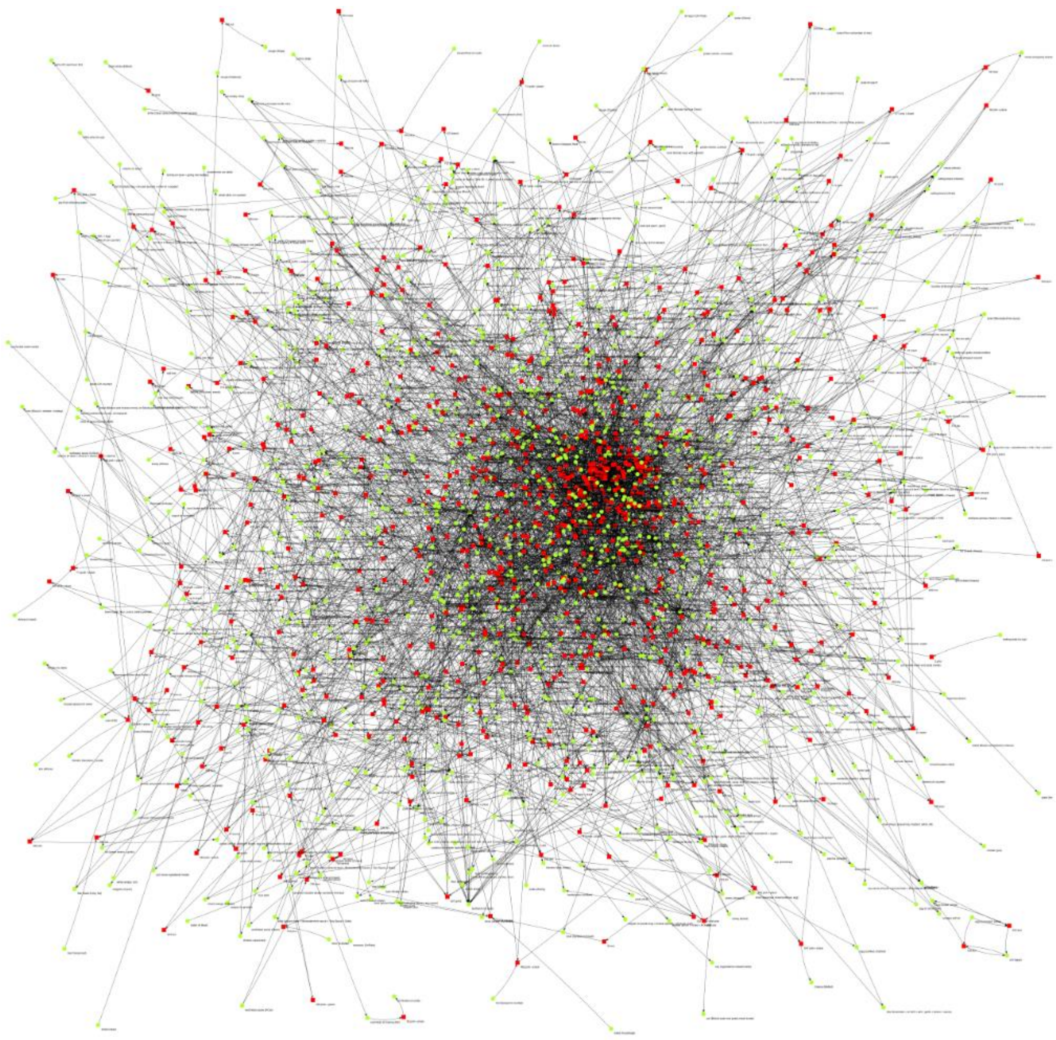

Figure 3: Our current universal FOON that is constructed from 56 online cooking videos.

Acknowledgements

This material is based upon work supported by the National Science Foundation under Grant No. 1421418.

References

E. Aksoy, A. Abramov, F. Worgotter, and B. Dellen. Categorizing object-action relations from semantic scene graphs. In *IEEE Intl. Conference on Robotics and Automation (ICRA)*, pages 398–405, 2010.

Dominik Jain, Lorenz Mosenlechner, and Michael Beetz. Equipping robot control programs with first-order probabilistic reasoning capabilities. In *IEEE International Conference on Robotics and Automation (ICRA)*, pages 3626–3631. IEEE, 2009.

G. D. Konidaris, S.R. Kuindersma, R.A. Grupen, and A.G Barto. Robot learning from demonstration by constructing skill trees. *Intl J Robotics Research*, 31(3):360–375, 2012.

B. D. Argall, S. Chernova, and et al. A survey of robot learning from demonstration. *Robotics and Autonomous Systems*, 57(5):469–483, 2009.

Y. Sun, S. Ren, and Y. Lin. Object-object interaction affordance learning. *Robotics and Autonomous Systems*, 2013.

Shaogang Ren and Yu Sun. Human-object-object-interaction affordance. In *Workshop on Robot Vision*, 2013.

M. E. J. Newman. *Networks: An Introduction*. Oxford University Press, USA, 2010.

Y. Huang and Y. Sun. Generating manipulation trajectories using motion harmonics. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 4949–4954, 2015.