Concluding Slides

Summary

- Overviewed the related cognitive and linguistic studies for representation of spatial and temporal meaning.
- Introduced the related resources and benchmarks for spatial and temporal information extraction and training models for spatiotemporal representations to help reasoning.
- Discussed mapping language to 1D/2D/3D/4D representations to n help human-like spatial and temporal reasoning.
- Evaluation of Large Language models of spatial temporal reasoning in multimodal setting.
- Spatial Commonsense reasoning
- We reviewed several downstream tasks such as situated grounding in multimodal setting, Navigation, Robot path finding, self-driving cars, human-object interaction, extraction of events timelines and more.

Summary

- Overviewed the related cognitive and linguistic studies for representation of spatial and temporal meaning.
- Introduced the related resources and benchmarks for spatial and temporal information extraction and training models for spatiotemporal representations to help reasoning.
- Discussed mapping language to 1D/2D/3D/4D representations to n help human-like spatial and temporal reasoning.
- Evaluation of Large Language models of spatial temporal reasoning in multimodal setting.
- Spatial Commonsense reasoning
- We reviewed several downstream tasks such as situated grounding in multimodal setting, Navigation, Robot path finding, self-driving cars, human-object interaction, extraction of events timelines and more.
    - This is not complete by any means…

Future Directions

- There is a gap between the past studies and what is used in current deep learning models for downstream tasks.
  - *The current deep architectures ignore the cognitive linguistic studies.*
- Still lack of benchmarks to evaluate the capabilities of the deep architectures and language models for spatial and temporal language understanding.
- Spatial language understanding needs commonsense about object affordances and real-world situations.
- We need more sophisticated models that take the explicit semantics into account to be able to rely on them in real-world scenarios and unobserved situations.
- Novel Pre-training and fine-tuning, data-augmentation, for symbolic conceptualizations
  - *Exploiting symbolic semantic abstraction for minimum task supervision for more generalizable models*
  - *Exploiting interactions for gaining spatial knowledge with indirect supervision*

A New Book Coming based on our Tutorials

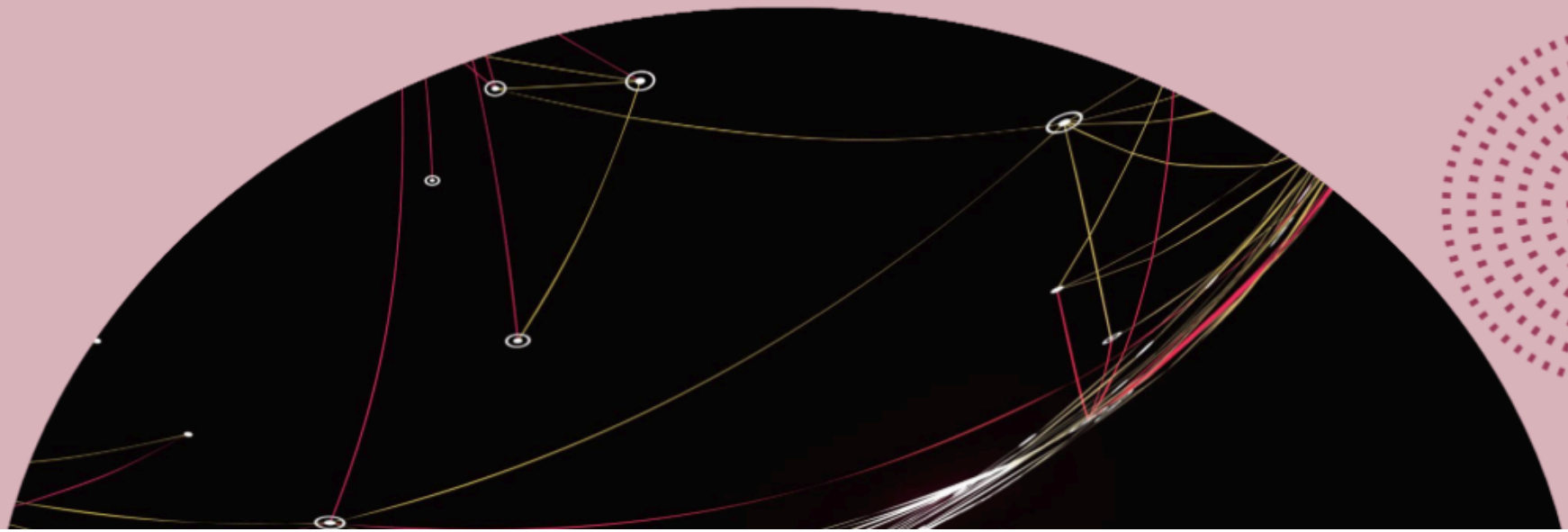
NEW BOOK: First Announcement

Spatial Language Understanding:
Representation, Reasoning, and Grounding

by Parisa Kordjamshidi, James Pustejovsky and Marie-Francine Moens

SYNTHESIS SERIES ON HUMAN LANGUAGE TECHNOLOGIES

Springer Nature



Tutorial Webpage and more …

We will add all the related info and link in here including the slides of past tutorials :

https://spatial-language-tutorial.github.io

Spatial Language Understanding (SpLU) and Robo-NLP workshop colocated with **ACL-2024** in Bangkok:

https://splu-robonlp-2024.github.io/

Final Word

We need more people work on this…

https://spatial-language-tutorial.github.io

MICHIGAN STATE
UNIVERSITY

European Research Council
Established by the European Commission

Final Word

We need more people work on this…

Fei-Fei Li:

With Spatial Intelligence, AI will Understand the Real World!

https://spatial-language-tutorial.github.io

aws

MICHIGAN STATE UNIVERSITY

KU LEUVEN

Brandeis

erc

European Research Council

Established by the European Commission

Final Word

Thank you! Questions?

We need more people work on this…

Fei-Fei Li:

With Spatial Intelligence, AI will Understand the Real World!

https://spatial-language-tutorial.github.io

MICHIGAN STATE
UNIVERSITY

European Research Council
Established by the European Commission

[Barclay, Michael & Galton, Antony. (2008). A Scene Corpus for Training and Testing Spatial Communication Systems.]

[Kordjamshidi, Roth, Moens,. BMC-BioInformatics. Structured learning for spatial information extraction from biomedical text: bacteria biotopes, 2015.]

[J. A. Bateman, J. Hois, R. Ross, and T. Tenbrink. A linguistic ontology of space for natural language processing. Artificial Intelligence, 174(14):1027–1071, 2010.]

[L. Talmy, The fundamental system of spatial schemas in language, in: From Perception to Meaning: Image Schemas in Cognitive Linguistics, Mouton de Gruyter, Berlin, 2006, pp. 37–47.]

[M. Bierwisch, How much space gets into language, in: P. Bloom, M.A. Peterson, L. Nadel, M.F. Garrett (Eds.), Language and Space, MIT Press, Cambridge,MA, 1999, pp. 31–76.]

[J. Bateman, T. Tenbrink, and S. Farrar. The role of conceptual and linguistic ontologies in discourse. Discourse Processes , 44(3):175–213, 2007.]

[J. Zlatev. Spatial semantics. In D. Geeraerts and H. Cuyckens, editors, The Oxford Handbook of Cognitive Linguistics , pages 318–350. Oxford Univ. Press, 2007.]

[Cohn, Anthony & Hazarika, Shyamanta. (2001). Qualitative Spatial Representation and Reasoning: An Overview. Fundamenta Informaticae, 46. 1-29]

[Andrew U. Frank, Qualitative Spatial Reasoning: Cardinal Directions as an Example, Geographical Information Systems 10(3):269-290, 1996]

[Max J. Egenhofer and Robert D. Franzosa, Point-set topological spatial relations, International Journal of Geographical Information Systems, 1991]

[Handbook of linguistic annotation, N Ide, J Pustejovsky, Springer, 2017.]

[Kordjamshidi et.al, 2012, Learning to interpret spatial natural language in terms of qualitative spatial relations Series Explorations in Language and Space.]

[P Kordjamshidi, M Van Otterlo, MF Moens, Spatial role labeling: Towards extraction of spatial relations from natural language ACM-Transactions in speech and language processing, 2011]

[P Kordjamshidi, P Frasconi, M Van Otterlo, MF Moens, L De Raedt, Relational learning for spatial relation extraction from natural language; International Conference on Inductive Logic Programming, ILP proceedings, LNCS, 2012]

[Kordjamshidi, P., van Otterlo, M., Moens, M.F.. From language towards formal spatial calculi. Computational Models of Spatial Language Interpretation Workshop (COSLI-2010) at COSIT.]

[Kordjamshidi, P., van Otterlo, M., Moens, M.F.. From language towards formal spatial calculi. Computational Models of Spatial Language Interpretation Workshop (COSLI-2010) at COSIT.]

[Kordjamshidi et al. SemEval2012]

[Kolomiyets, et.al. SemEval2013]

[Kordjamshidi, Moens. Global machine learning for spatial ontology population; Journal of Web Semantics, 2015]

[James Pustejovsky, Parisa Kordjamshidi, Marie-Francine Moens, Aaron Levine, Seth Dworman, Zachary Yocum. SemEval-2015 Task 8: SpaceEval; SemEval2015 workshop.]

[Soham Dan, Parisa Kordjamshidi, Julia Bonn, Archna Bhatia, Zheng Cai, Martha Palmer, Dan Roth, Soham Dan, Parisa Kordjamshidi, Julia Bonn, Archna Bhatia, Zheng Cai, Martha Palmer, Dan Roth. LREC 2020.]

[A dataset of chest X-ray reports annotated with Spatial Role Labeling annotations, Surabhi Datta, Kirk Roberts, In J Biomed Inform, 2020]
[Rad-SpatialNet: A Frame-based Resource for Fine-Grained Spatial Relations in Radiology Reports, Surabhi Datta, Morgan Ulinski, Jordan Godfrey-Stovall, Shekhar Khanpara, Roy F. Riascos-Castaneda, Kirk Roberts, LREC 2020.]

[SpatialNet: A Declarative Resource for Spatial Relations, Morgan Ulinski, Bob Coyne, Julia Hirschberg, SpLU-RoboNLP-2019]

[D. Schlangen, S. Zarrieß, and C. Kennington. Resolving references to objects in photographs using the words-as-classifiers model, ACL-2016.]

[Visually Guided Spatial Relation Extraction from Text, Rahgooy, Manzoor, Kordjamshidi, NAACL-2018]

[Manzoor, Kordjamshidi, Anaphora Resolution for Improving Spatial Relation Extraction from Text; NAACL, 2018, SpLU workshop]

[A Corpus for Reasoning about Natural Language Grounded in Photographs, Alane Suhr, Stephanie Zhou, Ally Zhang, Iris Zhang, Huajun Bai, Yoav Artzi, ACL-2019.]

[Making the V in VQA Matter: Elevating the Role of Image Understanding in Visual Question Answering, Yash Goyal and Tejas Khot and Douglas Summers-Stay and Dhruv Batra and Devi Parikh, CVPR-2017.]

[GQA: A New Dataset for Real-World Visual Reasoning and Compositional Question Answering, Hudson, Manning. CVPR-2019.]

[Chen Zheng, Quan Guo, Parisa Kordjamshidi. Cross-Modality Relevance for Reasoning on Language and Vision, Annual Conference of the Association for Computational Linguistics, ACL-2020.]

[SpaRTQA: A Textual Question Answering Benchmark for Spatial Reasoning. Roshanak Mirzaee, Hossein R. Faghihi, Qiang Ning and Parisa Kordjamshidi, EMNLP-2020, Spatial Language Understanding workshop, non-arxival.]

[Learning to Follow Navigational Directions , Adam Vogel and Dan Jurafsky, ACL-2010.]

[Toward Understanding Natural Language Directions, 5th ACM/IEEE International Conference on Human-Robot Interaction (HRI), Thomas Kollar, Stefanie Tellex, Deb Roy, Nicholas Roy, 2010]

[Anderson P, Wu Q, Teney D, et al. Vision-and-language navigation: Interpreting visually-grounded navigation instructions in real environments. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018: 3674-3683.]

[Are You Looking? Grounding to Multiple Modalities in Vision-and-Language Navigation, Ronghang Hu, Daniel Fried, Anna Rohrbach, Dan Klein, Trevor Darrell, Kate Saenko. ACL-2019.]

[Self-Monitoring Navigation Agent via Auxiliary Progress Estimation, Chih-Yao Ma, Jiasen Lu, Zuxuan Wu, Ghassan AlRegib, Zsolt Kira, Richard Socher, Caiming Xiong. ICLR-2019.]

[Learning to Navigate Unseen Environments: Back Translation with Environmental Dropout, Hao Tan, Licheng Yu, Mohit Bansal. NAACL-2019.]

[Vision-and-Language Navigation by Reasoning over Spatial Configurations. Yue Zhang, Quan Guo and Parisa Kordjamshidi, SpLU-2020 workshop at EMNLP, nonArxival.]

[J. Hois and O. Kutz. Natural language meets spatial calculi. In C. Freksa, N. S. Newcombe, P. Gärdenfors, and S. Wölfl, editors, Spatial Cognition VI. Learning, Reasoning, and Talking about Space , volume 5248 of LNCS, Springer, 2008.]

[Inderjeet Mani, et, al. (2009) SpatialML: Annotation Scheme, Resources, and Evaluation, MITRE Corporation.]

[Kordjamshidi, P., van Otterlo, M., Moens, M. F. (2010). Spatial role labeling: Task definition and annotation scheme. Proceedings of the Seventh conference on International Language Resources and Evaluation (LREC'10).]

[J. Pustejovsky and J. L. Moszkowicz. Integrating motion predicate classes with spatial and temporal annotations. In Donia Scott and Hans Uszkoreit, editors, COLING 2008: Companion volume D, Posters and Demonstrations , pages 95–98, 2008.]

[J. Pustejovsky and J. L. Moszkowicz. The role of model testing in standards development: The case of iso-space. In Proceedings of LREC'12 , pages 3060–3063. European Language Resources Association (ELRA), 2012.]

Jesse Thomason, Jivko Sinapov, Raymond Mooney, and Peter Stone. 2018. Guiding exploratory behaviors for multi-modal grounding of linguistic descriptions . In Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence (AAAI-18) .

S. Tellex, T. Kollar, S. Dickerson, M. R. Walter, A. G. Banerjee, S. Teller, and N. Roy. 2011. Understanding natural language commands for robotic navigation and mobile manipulation. In Proceedings of the National Conference on Artificial Intelligence (AAAI) .

Alane Suhr, Stephanie Zhou, Iris Zhang, Huajun Bai, and Yoav Artzi. 2018. A corpus for reasoning about natural language grounded in photographs. arXiv preprint arXiv:1811.00491 .

Kristin Stock, Robert C. Pasley, Zoe Gardner, Paul Brindley, Jeremy Morley, and Claudia Cialone. 2013. Creating a corpus of geospatial natural language. In Spatial Information Theory , pages 279–298, Cham. Springer International Publishing.

Gengchen Mai, Krzysztof Janowicz, Ling Cai, Rui Zhu, Blake Regalia, Bo Yan, Meilin Shi, and Ni Lao. 2020. Se-kge: A location-aware knowledge graph embedding model for geographic question answering and spatial semantic lifting . Transactions in GIS , 24(3):623–655.

Yunchao Liu, Jiajun Wu, Zheng Wu, Daniel Ritchie, William T. Freeman, and Joshua B. Tenenbaum. 2019. Learning to describe scenes with programs. In International Conference on Learning Representations .

Nikhil Krishnaswamy, Scott Friedman, and James Pustejovsky. 2019. Combining deep learning and qualitative spatial reasoning to learn complex structures from sparse examples with noise. In Proceedings of the thirty-third AAAI Conference on Artificial Intelligence .

John D. Kelleher and Geert-Jan M. Kruijff. 2006. Incremental generation of spatial referring expressions in situated dialog . In Proceedings of the 21st International Conference on Computational Linguistics and 44th Annual Meeting of the Association for Computational Linguistics, ages 1041–1048, Sydney, Australia. Association for Computational Linguistics.

Guillem Collell Talleda and Marie-Francine Moens. 2018. Learning representations specialized in spatial knowledge: Leveraging language and vision . Transactions of the Association for Computational Linguistics , 6:133–144.