A Model for Attention-Driven Judgements in Type Theory with Records

Simon Dobnik[1] and John D. Kelleher[2]
[1]University of Gothenburg and [2]Dublin Institute of Technology

CLASP research seminar
12 May 2016

A knowledge representation framework for robots

- Information fusion: interactively combining, comparing and reasoning with information from perceptual and conceptual domains

A knowledge representation framework for robots

- Information fusion: interactively combining, comparing and reasoning with information from perceptual and conceptual domains
- Layered approach (Kruijff et al., 2007)
  - meaning representations are modular and independent
  - interfaces to mediate between the levels

A knowledge representation framework for robots

UNIVERSITY OF GOTHENBURG

- Information fusion: interactively combining, comparing and reasoning with information from perceptual and conceptual domains
- Layered approach (Kruijff et al., 2007)
  - meaning representations are modular and independent
  - interfaces to mediate between the levels
- TTR (Cooper, 2012; Cooper et al., 2014): from types of sensory readings to types of information states (ISs) (Dobnik et al., 2014)

Why TTR?

UNIVERSITY OF GOTHENBURG

- Agent centred: an agent makes judgements that an object a is of type T or $a : T$
- Type system is learned as agent interacts with its environment (perception, Dobnik et al. (2013)) and other agents (dialogue, (Larsson, 2013))
- Open to revision as agent enters new environments or dialogue contexts
- Convergence of the type system across agents is ensured by being located in the same perceptual and linguistic contexts
- A different view in computational linguistics but a standard view in mobile robotics (Dissanayake et al., 2001)

The flexibility of the type system

- Rich type system:

$$\left[\begin{array}{lcl} \text{a} & = & \text{ind}_{26} \\ \text{sr} & = & [[34,24],[56,78]\ldots] \\ \text{loc} & = & [45,78,0.34] \end{array}\right] : \left[\begin{array}{lcl} \text{a} & : & \text{Ind} \\ \text{sr} & : & \text{list(list(Real))} \\ \text{loc} & : & \text{list(Real)} \end{array}\right]$$

The flexibility of the type system

- Rich type system:

$$\begin{bmatrix} \text{a} & = & \text{ind}_{26} \\ \text{sr} & = & [[34,24],[56,78]\ldots] \\ \text{loc} & = & [45,78,0.34] \end{bmatrix} : \begin{bmatrix} \text{a} & : & \text{Ind} \\ \text{sr} & : & \text{list(list(Real))} \\ \text{loc} & : & \text{list(Real)} \end{bmatrix}$$

- Intensionality: a situation may be assigned more than one record type *Left*, *Near*, *At*, *Behind*

The flexibility of the type system

- Rich type system:

$$\begin{bmatrix} \text{a} & = & \text{ind}_{26} \\ \text{sr} & = & [[34,24],[56,78]\ldots] \\ \text{loc} & = & [45,78,0.34] \end{bmatrix} : \begin{bmatrix} \text{a} & : & \text{Ind} \\ \text{sr} & : & \text{list(list(Real))} \\ \text{loc} & : & \text{list(Real)} \end{bmatrix}$$

- Intensionality: a situation may be assigned more than one record type *Left*, *Near*, *At*, *Behind*
- Sub-typing: $\textit{Chair} \sqsubseteq \textit{Object} \sqsubseteq \textit{Physical entity} \sqsubseteq \textit{Entity}$ and if $s : \textit{Chair}$ then $s : \textit{Object}$, $s : \textit{Physical entity}$ and $s : \textit{Entity}$

The flexibility of the type system

- Rich type system:
$$\left[\begin{array}{lll} \text{a} & = & \text{ind}_{26} \\ \text{sr} & = & [[34,24],[56,78]\ldots] \\ \text{loc} & = & [45,78,0.34] \end{array}\right] : \left[\begin{array}{lll} \text{a} & : & \text{Ind} \\ \text{sr} & : & \text{list(list(Real))} \\ \text{loc} & : & \text{list(Real)} \end{array}\right]$$
- Intensionality: a situation may be assigned more than one record type *Left*, *Near*, *At*, *Behind*
- Sub-typing: $\textit{Chair} \sqsubseteq \textit{Object} \sqsubseteq \textit{Physical entity} \sqsubseteq \textit{Entity}$ and if $s : \textit{Chair}$ then $s : \textit{Object}$, $s : \textit{Physical entity}$ and $s : \textit{Entity}$
- Component types: $s : \textit{Left}$ and $s : \textit{Table-Left-Chair}$

The flexibility of the type system

UNIVERSITY OF GOTHENBURG

- Rich type system:

$$\left[\begin{array}{lcl} \text{a} & = & \text{ind}_{26} \\ \text{sr} & = & [[34,24],[56,78]\ldots] \\ \text{loc} & = & [45,78,0.34] \end{array}\right] : \left[\begin{array}{lcl} \text{a} & : & \text{Ind} \\ \text{sr} & : & \text{list(list(Real))} \\ \text{loc} & : & \text{list(Real)} \end{array}\right]$$

- Intensionality: a situation may be assigned more than one record type *Left*, *Near*, *At*, *Behind*
- Sub-typing: *Chair* $\sqsubseteq$ *Object* $\sqsubseteq$ *Physical entity* $\sqsubseteq$ *Entity* and if s : *Chair* then s : *Object*, s : *Physical entity* and s : *Entity*
- Component types: s : *Left* and s : *Table-Left-Chair*
- Dependent types: s : *Table-Left-Chair* and s : *Table* and s : *Left*

The computational cost of type judgements

- Each type assignment involves a binary judgement

The computational cost of type judgements

- Each type assignment involves a binary judgement
- An agent with n types can make n judgements of each situation

The computational cost of type judgements

- Each type assignment involves a binary judgement
- An agent with n types can make n judgements of each situation
- A learning agent is faced with 2^n possible assignments

The computational cost of type judgements

- Each type assignment involves a binary judgement
- An agent with n types can make n judgements of each situation
- A learning agent is faced with 2^n possible assignments

The computational cost of type judgements

- Each type assignment involves a binary judgement
- An agent with n types can make n judgements of each situation
- A learning agent is faced with 2^n possible assignments
  $n = 3$, $2^3 = 8$: $\{\}$, $\{T_1\}$, $\{T_2\}$, $\{T_3\}$, $\{T_1, T_2\}$, $\{T_1, T_3\}$, $\{T_2, T_3\}$ and $\{T_1, T_2, T_3\}$

Sub-typing as a lattice

- Sub-typing can be inferred by comparing record structures
- (Hough and Purver, 2014) organise types in a lattice by subtype relation for incremental inference
- Allows us to prune sub-types of an incompatible type
- Taxonomic or categorical relations
- Do humans judge situations from most general to most specific?

Thematic relations

UNIVERSITY OF GOTHENBURG

- Require priming what to expect in the current state given the knowledge about the world.
- Thematic relations: spatial, temporal, causal or functional relations between individuals occuring in the same situations Lin and Murphy (2001); Estes et al. (2011)
- Type resources (Cooper, 2008) that are employed and learned in different situational contexts

In this talk

UNIVERSITY OF GOTHENBURG

- What drives the creation of resources/thematic relations?
- How are bundles of types selected and primed in particular situational contexts?
- How can we model them computationally - for an application of TTR in a situated robot?

Attention

UNIVERSITY OF GOTHENBURG

- Bottom-up attention: not under conscious control and task independent

Attention

- Bottom-up attention: not under conscious control and task independent
- Top-down attention: consciously directed and task dependent/primed

Attention

- Bottom-up attention: not under conscious control and task independent
- Top-down attention: consciously directed and task dependent/primed
- A shared resource that can spread across multiple tasks to different degrees depending on the difficulty of the task and attention policy (Kahneman, 1973)

Attention

- Bottom-up attention: not under conscious control and task independent
- Top-down attention: consciously directed and task dependent/primed
- A shared resource that can spread across multiple tasks to different degrees depending on the difficulty of the task and attention policy (Kahneman, 1973)
- What are the conditions under which the perception of task irrelevant distractors is prevented and at what stage?

Attention driven judgements

- Load Theory (LT) (Lavie et al., 2004):
  - Perceptional selection: the more perceptual load the less capacity to perceive distractor objects
  - Cognitive control: active processing prioritisation of task-relevant stimuli and reduction of perceived distractors

Attention driven judgements

UNIVERSITY OF GOTHENBURG

- Load Theory (LT) (Lavie et al., 2004):
  - Perceptional selection: the more perceptual load the less capacity to perceive distractor objects
  - Cognitive control: active processing prioritisation of task-relevant stimuli and reduction of perceived distractors
- Attention driven judgements:
  - Pre-attentive
  - Task and context induced

Pre-attentive judgements

- Perceptional selection mechanism of LT
- Iconic representations Harnad (1990)
- Ullman (1984): basic representations of visual environment and visual routines
- Segmentation of a visual scene into entities and background

Pre-attentive judgements

UNIVERSITY OF GOTHENBURG

- Perceptional selection mechanism of LT
- Iconic representations Harnad (1990)
- Ullman (1984): basic representations of visual environment and visual routines
- Segmentation of a visual scene into entities and background
- Linked to agent's biology and embodiment, sensors and actuators: finite in number and "basic"
- Fundamental to the agent's basic operation: made continuously
- The judgements are pushed to the IS at a rate determined by LT

Task-induced and context-induced judgements

- Cognitive control mechanism of LT
- An agent is making a cup of tea in the kitchen on the second floor at FLOV
- Task-induced judgements: primed by a default set of objects and actions associated with the task
- Context-induced judgements: primed by the context in which the task is taking place
- Both kinds of judgements interact
- Making a judgement belonging to a task or a context primes the agent to further judgements of that task or context

Control mechanism for TI and CI judgements

- Organisation of agent's type inventory in memory according to thematic relations (Lin and Murphy, 2001; Estes et al., 2011)
- How is type inventory organised this way used in making primed typed judgements following the LT?

Cognitive states

- An agent experiences the world through perception, embodiment and linguistic interaction
- Experiencing tasks and contexts, an agent forms associations between types co-occuring in its memory
- Associations clusters are modelled as cognitive states
- An agent is not conscious of its states
- ... but they prime the agent to particular types of situations
- An agent may be in one or more state at the same time
- A particular type may be associated with more than one state

Cognitive states, II

UNIVERSITY OF GOTHENBURG

Relations between states are computationally more tractable than relations between types

- States are built bottom up as agent discovers new situations
- Constrained by the environment in which it operates
- Can only discover a finite set of states in its lifetime
- Has a strong learning bias for making generalisations

Cognitive states: computational model

- Latent Dirichlet Allocation (LDA) Blei et al. (2003): document.word := topic and memory.type := state
- Hierarchical Dirichlet Process Teh et al. (2006) for unknown number of topics/states

Decision about states

We need...

- Update mechanism for the posterior distribution over states
- A decision mechanism regarding which types to be primed to based on the posterior probability distribution over states

Posterior probability over states

- Probability of states at $t-1$
- The task and context judgements the agent has made following the priming at $t-n$, n is the length of history
- Pre-attentive judgements at t reflecting perceptual change in its world

$$
\begin{aligned}
&P(s_t|Pre_t, Task_{t-1}, Cont_{t-1}, AS_{t-1}) = \\
&\quad \eta \times P(Pre_t, Task_{t-1}, Cont_{t-1}, AS_{t-1}|s_t) \\
&\quad \times P(s_t)
\end{aligned}
$$

Posterior probability over states, II

- Assuming conditional independence:

$$
\begin{aligned}
&P(s_t|Pre_t, Task_{t-1}, Cont_{t-1}, AS_{t-1}) = \\
&\quad \eta \times P(Pre_t|s_t) \times P(Task_{t-1}|s_t) \\
&\quad \times P(Cont_{t-1}|s_t) \times P(AS_{t-1}|s_t) \times P(s_t)
\end{aligned}
$$

Priming of types, I

- Select $s_t \in S$ with the maximum a posteriori probability
- Load the types from s_t into short-term memory
- $+$ and –:
  - $+$ simple
  - – agent assumes it is only in 1 state
  - – may end up switching between two states

Priming of types, II

- Rank and prune the state set using the posterior probability: active relevant states (AS)
- The threshold determined by available resources: perceptual selection and cognitive control
- Renormalise the probability distribution over AS
- Compute a posterior probability over the set of types in AS using a Bayes optimal classifier
- Using the posterior probability over types, rank and prune the set of types
- Load the set of unpruned types into working memory

Probability of a type

$$P(T_{i,t}|Pre_t, Task_{t-1}, Cont_{t-1}, AS_{t-1}) = \sum_{s_i \in AS_t} P(T_i|s_i) \times P(s_i|Pre_t, Task_{t-1}, Cont_{t-1}, AS_{t-1})$$

Probability of a type

$$P(T_{i,t}|Pre_t, Task_{t-1}, Cont_{t-1}, AS_{t-1}) = \sum_{s_i \in AS_t} P(T_i|s_i) \times P(s_i|Pre_t, Task_{t-1}, Cont_{t-1}, AS_{t-1})$$

+ A type is associated with more than one state

Probability of a type

$$P(T_{i,t}|Pre_t, Task_{t-1}, Cont_{t-1}, AS_{t-1}) = \sum_{s_i \in AS_t} P(T_i|s_i) \times P(s_i|Pre_t, Task_{t-1}, Cont_{t-1}, AS_{t-1})$$

- A type is associated with more than one state
- More than one state may be active at one time

Probability of a type

$$P(T_{i,t}|Pre_t, Task_{t-1}, Cont_{t-1}, AS_{t-1}) = \sum_{s_i \in AS_t} P(T_i|s_i) \times P(s_i|Pre_t, Task_{t-1}, Cont_{t-1}, AS_{t-1})$$

- A type is associated with more than one state
- More than one state may be active at one time
- Types and states as associated probabilistically: $P(T_i|s_i)$

Probability of a type

$$P(T_{i,t}|Pre_t, Task_{t-1}, Cont_{t-1}, AS_{t-1}) = \sum_{s_i \in AS_t} P(T_i|s_i) \times P(s_i|Pre_t, Task_{t-1}, Cont_{t-1}, AS_{t-1})$$

- A type is associated with more than one state
- More than one state may be active at one time
- Types and states as associated probabilistically: $P(T_i|s_i)$
- Several states may be maximising a probability of a particular type

Probability of a type

UNIVERSITY OF GOTHENBURG

$$P(T_{i,t}|Pre_t, Task_{t-1}, Cont_{t-1}, AS_{t-1}) = \sum_{s_i \in AS_t} P(T_i|s_i) \times P(s_i|Pre_t, Task_{t-1}, Cont_{t-1}, AS_{t-1})$$

+ A type is associated with more than one state
+ More than one state may be active at one time
+ Types and states as associated probabilistically: $P(T_i|s_i)$
+ Several states may be maximising a probability of a particular type
+ The system is more stable in making decisions than argmax

Probability of a type, II

UNIVERSITY OF GOTHENBURG

+ The more judgements we make the more we reduce the ambiguity of being in several states.
– Calculating posterior probabilities of types in active states is computationally more expensive than calculating probabilities of states
+ The aggressiveness of the pruning criteria: Load Theory

Conclusions and future work

- Attention-driven type judgements in an interacting agent inspired by discovery of thematic relations and sharing of cognitive resources

Conclusions and future work

UNIVERSITY OF GOTHENBURG

- Attention-driven type judgements in an interacting agent inspired by discovery of thematic relations and sharing of cognitive resources
- Agent maintains:
  - a distribution set of cognitive states
  - a distribution over set of types in the active states
  - the number of active states is controlled by available cognitive resources

Conclusions and future work

- Attention-driven type judgements in an interacting agent inspired by discovery of thematic relations and sharing of cognitive resources
- Agent maintains:
  - a distribution set of cognitive states
  - a distribution over set of types in the active states
  - the number of active states is controlled by available cognitive resources
- A general problem for agent making classification:
  - visual search in robotics (Sjöö, 2011; Kunze et al., 2014)
  - situated dialogue: disambiguation of speakers utterances/topic priming
  - situated dialogue: generating new utterances/topic modelling

References I

Blei, D. M., A. Y. Ng, and M. I. Jordan (2003). Latent dirichlet allocation. *Journal of Machine Learning research (JMLR) 3*, 993–1022.

Cooper, R. (2008). Type theory with records and unification-based grammar. In F. Hamm and S. Kepser (Eds.), *Logics for Linguistic Structures. Festschrift for Uwe Mönnich*, Volume 201, pp. 9. Walter de Gruyter.

Cooper, R. (2012). Type theory and semantics in flux. In R. Kempson, N. Asher, and T. Fernando (Eds.), *Handbook of the Philosophy of Science*, Volume 14 of *General editors: Dov M Gabbay, Paul Thagard and John Woods*. Elsevier BV.

Cooper, R., S. Dobnik, S. Lappin, and S. Larsson (2014, 27 April). A probabilistic rich type theory for semantic interpretation. In R. Cooper, S. Dobnik, S. Lappin, and S. Larsson (Eds.), *Proceedings of the EACL 2014 Workshop on Type Theory and Natural Language Semantics (TTNLS)*, Gothenburg, Sweden, pp. 72–79. Association for Computational Linguistics.

References II

Dissanayake, M. W. M. G., P. M. Newman, H. F. Durrant-Whyte, S. Clark, and M. Csorba (2001). A solution to the simultaneous localization and map building (SLAM) problem. *IEEE Transactions on Robotic and Automation 17*(3), 229–241.

Dobnik, S., R. Cooper, and S. Larsson (2013). Modelling language, action, and perception in Type Theory with Records. In D. Duchier and Y. Parmentier (Eds.), *Constraint Solving and Language Processing: 7th International Workshop, CSLP 2012, Orléans, France, September 13–14, 2012, Revised Selected Papers*, Volume 8114 of *Lecture Notes in Computer Science*, pp. 70–91. Springer Berlin Heidelberg.

Dobnik, S., R. Cooper, and S. Larsson (2014, 13 November). Type Theory with Records: a general framework for modelling spatial language. In S. Dobnik, R. Cooper, and S. Larsson (Eds.), *Proceedings of The Second Workshop on Action, Perception and Language (APL'2)*, The Fifth Swedish Language Technology Conference (SLTC), Uppsala, Sweden.

References III

Estes, Z., S. Golonka, and L. L. Jones (2011). Thematic thinking: The apprehension and consequences of thematic relations. In B. Ross (Ed.), *The Psychology of Learning and Motivation*, Volume 54, pp. 249–294. Burlington: Academic Press.

Harnad, S. (1990, June). The symbol grounding problem. *Physica D 42*(1–3), 335–346.

Hough, J. and M. Purver (2014, April). Probabilistic type theory for incremental dialogue processing. In *Proceedings of the EACL 2014 Workshop on Type Theory and Natural Language Semantics (TTNLS)*, Gothenburg, Sweden, pp. 80–88. Association for Computational Linguistics.

Kahneman, D. (1973). *Attention and effort*. Englewood Cliffs, N.J.: Prentice-Hall.

Kruijff, G.-J. M., H. Zender, P. Jensfelt, and H. I. Christensen (2007). Situated dialogue and spatial organization: what, where... and why? *International Journal of Advanced Robotic Systems 4*(1), 125–138. Special issue on human and robot interactive communication.

References IV

UNIVERSITY OF GOTHENBURG

Kunze, L., C. Burbridge, and N. Hawes (2014, March, 24–26). Bootstrapping probabilistic models of qualitative spatial relations for active visual object search. In *AAAI Spring Symposium 2014 on Qualitative Representations for Robots*, Stanford University in Palo Alto, California, US.

Larsson, S. (2013, December 18). Formal semantics for perceptual classification. *Journal of Logic and Computation online*, 1–35.

Lavie, N., A. Hirst, J. W. de Fockert, and E. Viding (2004). Load theory of selective attention and cognitive control. *Journal of Experimental Psychology: General 133*(3), 339–354.

Lin, E. L. and G. L. Murphy (2001). Thematic relations in adults' concepts. *Journal of experimental psychology: General 130*(1), 3–28.

Sjöö, K. (2011). *Functional understanding of space: Representing spatial knowledge using concepts grounded in an agent's purpose*. Ph. D. thesis, KTH, Computer Vision and Active Perception (CVAP), Centre for Autonomous Systems (CAS), Stockholm, Sweden.

References V

Teh, Y. W., M. I. Jordan, M. J. Beal, and D. M. Blei (2006). Hierarchical dirichlet processes. *Journal of the American Statistical Association 101*(476), 1566–1581.

Ullman, S. (1984). Visual routines. *Cognition 18*(1–3), 97–159.