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

Simon Dobnik
CLASP
University of Gothenburg, Sweden
`simon.dobnik@gu.se`

John D. Kelleher[*]
School of Computing
Dublin Institute of Technology, Ireland
`john.d.kelleher@dit.ie`

Abstract

This paper makes three contributions to the discussion on the applicability of Type Theory with Records (TTR) to embodied dialogue agents. First, it highlights the problem of type assignment or judgements in practical implementations which is resource intensive. Second, it presents a judgement control mechanism, which consists of grouping of types into clusters or states by their thematic relations and selection of types following two mechanisms inspired by the Load Theory of selective attention and cognitive control (Lavie et al., 2004), that addresses this problem. Third, it presents a computational framework, based on Bayesian inference, that offers a basis for future practical experimentation on the feasibility of the proposed approach.

1   Type Theory with Records

One of the central challenges for multi-modal dialogue systems is information fusion or how such a system can represent information from different domains, compare it, compose it, and reason about it. Typically, a situated agent will have to deal with information that comes from its perceptual sensors and will be represented as real-valued vectors and conceptual categories (some of which correspond to words in language) that are formed through cognitive processes in the brain. When situated agents are implemented practically one typically adopts a layered approach starting at the scene geometry and finishing at the level of the agent's knowledge about the objects and their interactions (Kruijff et al., 2006). Although, this approach may be good for practical reasons, for example there are pre-existing systems which may be organised in a

pipeline this way, this also assumes that representations and operations are distinct at each level and one needs to design interfaces that would mediate between these levels.

Type Theory with Records (TTR) (Cooper, 2005; Cooper, 2012; Cooper et al., 2015) provides a theory of natural language semantics which views meaning and reference assignment being in the domain of an individual agent who can make *judgements* about situations (or invariances in the world) of being of types (written as $a : T$). The type inventory of an agent is not static but is continuously refined through agent's interaction with its physical environment and with other agents through dialogue interaction which provides instances and feedback on what strategies to adopt to learn from these instances. The reason why agent's meaning representations or type inventories converge to an approximately identical inventory is that agents are situated in the identical or sufficiently similar physical environment and have grounded conversations with other agents; see for example the work of (Steels and Belpaeme, 2005) and (Larsson, 2013) for an approach in TTR. Having the capability to adjust the type representations they can adapt to new physical environments and new conversational exchanges. Such view is not novel to mobile robotics (Dissanayake et al., 2001) nor to approaches to semantic and pragmatics of dialogue (Clark, 1996), but it is novel to formal semantics (Dowty et al., 1981; Blackburn and Bos, 2005) which represents important body of work on how meaning is constructed compositionally and reasoned about. Overall, we see TTR as a highly fitting framework for modelling cognitive situated agents as it connects perception and high level semantics of natural language and vice versa.

The type system in TTR is rich in comparison to that found in traditional formal semantics (entities, truth values and function types constructed from these and other function types). In addition types

---

[*] Both authors contributed equally.

are used to model meaning in a proof-theoretic way rather than constraining model theoretic interpretation. Types in TTR can be either basic types such as *Ind* or *Real* or record types. Record types are represented as matrices containing label-value pairs where labels are constants and values can be either basic types, ptypes which act as type constructors and record types. The corresponding proof-objects of record types are records. These may be thought of as *iconic representations* of (Harnad, 1990) or sensory readings that an agent perceives as *sensory projections* of *objects* or *situations* in the world. The example below shows a judgement that a record (a matrix with = as a delimiter) containing a sensory reading is of a type (with : as a delimiter). The traditional distinction between symbolic and sub-symbolic knowledge is not maintained in this framework as both can be assigned appropriate types.

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

An important notion of TTR is that types are *intensional* which means that a given situation in the world may be assigned more than one type. For example, a sensory reading of a particular situation in the world involving spatial arrangement of objects captured as records of types shown in the previous example may be assigned several record types of spatial relations simultaneously, for example *Left*, *Near*, *At*, *Behind*, etc. Another important notion of TTR is *sub-typing* which allows comparison of types. In addition to the type in the previous example, let's call it T_1, the record can also be assigned the following two types T_2 and T_3 whereby the following sub-typing relation between them holds: $T_1 \sqsubseteq T_2 \sqsubseteq T_3$ where $\sqsubseteq$ reads as "is a subtype of".

$$T_2 = \begin{bmatrix} a & : & Ind \\ sr & : & list(list(Real)) \end{bmatrix} \\ T_3 = \begin{bmatrix} a & : & Ind \end{bmatrix}$$

Thirdly, types may be component types of other types, for example p(redicate)-type list(Real) is a component of the larger record type shown above. Finally, record types may also be dependent on other record types. A record type representing a geometric relation between two objects, for example *Left* is dependent on at least two record types shown in the previous example representing perceptual objects. The notion of dependent types is stronger than that of component types and is related to representing linguistic meaning. With missing information matching a component type an agent could still judge (with some error) that a situation is of that type whereas a judgement without a previous judgement of the dependent type would be impossible.

The rich type system of TTR and the relations between types give us a lot of flexibility in modelling natural language semantics in embodied dialogue agents. However, one practical problem that an application of TTR faces is how such an agent will cope with the an increasing number of types that it continuously acquires through learning and assign them effectively to every new situation it encounters given that such agent has limited processing resources. Since each type assignment involves a judgement (a probabilistic belief that something is of a type T) for each record of a situation an agent having an inventory of n types would make n judgements, a large proportion of which would yield very low or even zero probabilities as they will be irrelevant or very-little relevant for the current perceptual and conversational contexts. This is because due to the regularities in the world certain types would never be assigned or are very unlikely to be assigned in certain contexts.

(Hough and Purver, 2014) present a model where types are ordered in a lattice by $\sqsubseteq$ which drives incremental type checking for the purposes of resolution of incremental linguistic input or output which in itself is a different task to ours. The approach captures *taxonomic or categorial relations* encoded in types. As humans do not necessarily judge situations from most general to most specific or vice versa, the benefit of reducing judgements following taxonomic organisation of types would vary depending on the situation judged. Such knowledge would allow exclusion of judgements of sub-types of an incompatible type but agent's judgements could be further reduced if it were primed what to expect in its current state and its perceptual and conversational contexts by its knowledge about the world and the linguistic behaviour of its interlocutor captured in a model of *thematic relations*, that is spatial, temporal, causal or functional relations between individuals occuring in the same situations (Lin and Murphy, 2001;

Estes et al., 2011). Similarly, (Cooper, 2008) argue that agents organise their type inventory into resources that are employed and modified in different activities. If this is so, in addition to a reasoning mechanism on subtype relations humans must also rely on processes by which bundles of types are primed for in particular situational contexts. As a consequence agents will not need to check each situation (sensory reading in the form of a record) for every type in their inventory but only those that they are primed for. A property that such priming mechanism must take into account is that the more accurately an agent is primed by its contexts, the lower the uncertainty and hence the smaller the set of the types it is primed for.

In this paper we focus on the mechanisms that drive the discovery of thematic relations and propose a computational model how such relations are applied in interaction to prime an agent. The basic premise of the paper is that the mechanisms underpinning attention are fundamental to the control and priming of judgements in TTR. In Section 2, we introduce Load Theory of Attention. In Section 3, we present an account, based on Load Theory of Attention, of how two different kinds of TTR judgements can be controlled and primed in an agent. In Section 4, we introduce a mathematical framework that illustrates how an agent can maintain probability distributions over its cognitive states and types and use them in the priming process. In Section 5 we give a worked example of this framework priming an agent for judgements. Section 6 gives some remarks about its usability and presents our future work.

2 Load Theory of Attention

One of the major contributions of 20th century psychology has been the study and improved understanding of perceptual attention in humans. There is more than one type of attention mechanism. In particular, a distinction can be made between bottom-up attention and top-down attention processes. Bottom-up attention is automatic, task independent, not under conscious control and is attracted towards salient entities in the environment (e.g., moving object, singleton red objects, etc.). Top-down attention can be consciously directed by an agent and is dependent on the task they are carrying out as tasks will have different complexities. Sometimes top-down attention is described in terms of an agent being primed to respond to a

mental-set of perceptual stimuli that are relevant to the task they are consciously carrying out.

Early research on attention was based on the concept of a structural single channel bottleneck in perceptual processing (Broadbent, 1958; Welford, 1967). The early orthodoxy of attention as a bottleneck within a single channel has been challenged by several researchers (e.g., (Allport et al., 1972)) and more recent models have viewed attention as a shared resource or capacity that can be spread across multiple tasks simultaneously. For example, in the (Kahneman, 1973) theory of attention and effort the attention capacity can be focused on an individual task or shared across multiple tasks and the more difficult a task is the more attention is required by that task. Furthermore, the allocation of attention across tasks can be flexibly updated as the agent changes their attention policy from one moment to the next.

An enduring question within attention research has been to understand the conditions under which the perception of task irrelevant distractors is prevented. Most of this research in the 60s, 70s, and 80s was framed in terms of the early-late debate which focused on whether the structural bottleneck that excluded distractors occurred early or late in perceptual processing. Some researchers argued that attention could exclude early perceptual processing of distractors (e.g., (Treisman, 1969)) while others argued that distractor objects were perceptually processed and attention only affected post perceptual processing – such as working memory and response selection (e.g., (Duncan, 1980)). The reason for such a protracted debate was that there was a lot of evidence to support both views. Results from some studies indicated that unattended information went unnoticed (supporting an early filter) and other studies indicated that distractors were perceptually processed and interfered with task response (supporting a late filter).

A well regarded recent model of attention is perceptual load theory (Lavie et al., 2004). The concept of perceptual load is difficult to define but can be characterised in terms of the number of items that are perceptually available (the more items, the higher the load) and the demands of the perceptual task (e.g., selecting an object based on type and colour is more demanding then selecting an object based solely on type). Perceptual load also involves defining what constitutes an item in a display: (Lavie et al., 2004) give the example that

a string of letters can be considered one item (a word) or several items (letters) depending on the task. Perceptual load theory attempts to resolve the early-late debate using a model of attention that distinguishes between two mechanisms of selective attention: *perceptual selection* and *cognitive control*. Perceptual selection is a mechanism that excludes the perception of task irrelevant distractors under situations of high perceptual load; however, in situations of low perceptual load any spare capacity will spill over to the perception of distractor objects. The cognitive control mechanism is an active process that reduces the interference from perceived distractors on task response. It does so by actively maintaining the processing prioritisation of task relevant stimuli within the set of perceived stimuli.

3   Load Theory and type judgements

Agents learn types all the time by making generalisations of invariances in the world and information communicated to them through conversation (direct transferal of knowledge). However, in order to access the knowledge quickly and efficiently, they organise it in a certain way in memory. We propose a method of how an agent (i) organises its type inventory in memory and (ii) applies this type inventory using a model of attention that avoids the exponential problem of judgements it would have to make without prioritising its type checking. We turn to the second notion first, the priming of type judgements using a model of attention. Within this attention based account a distinction can be made between two types of judgements: (i) pre-attentive and (ii) task induced and context induced judgements.

3.1   Attention-driven judgements

Pre-attentive judgements are controlled by the perceptual selection mechanism of Load Theory. The result of a pre-attentive judgement is the introduction of a type into the working memory or information state in a dialogue model (Ginzburg and Fernández, 2010). Basic representations of visual environment (Ullman, 1984) such as segmentation of a visual scene into entities and background is an example of a pre-attentive judgement. At the very basic level these will be the iconic representations captured by agent's sensors (Harnad, 1990). *Task induced* and *context induced judgements* require conscious attention. As such, they are controlled

by the cognitive control mechanisms of Load Theory. These judgements are applied to types that are in working memory and result in new types being introduced to working memory. Task induced and context induced judgements are primed by the types associated via memory with the current activities that the agent is currently engaged in and their physical location. For example, if an agent is making a cup of tea there are a default set of objects relevant to that task that the agent will carry out a visual search for and purposefully recognise (the kettle, tea bags, cups, etc.). The definition of a set of relevant types corresponding to these objects can be understood as priming a set of task induced judgements related to the recognition of these objects. Finally, context induced judgements can be understood as task related judgements that are not by default related with the task but that are extensions to this set and are caused by the agent's interactions with other agents and the physical context of the task. For example, while an agent is making a cup of tea another agent warns them to take care because the plate beside the kettle is very hot or the agent may inadvertently touch the plate and sense the heat on its own. The judgement relating to the interpretation of the utterance "the plate beside the kettle" or the sensing of and predicting the desired reaction to a hot surface can be understood as a context induced judgement. The utterance or the hot plate is not a part of the task but is introduced in the context in which the task is taking place.

This raises a question of what mechanisms define these classes of type judgements. Pre-attentive type judgements are the judgements that are fundamental to the agent's basic operation and the agent is continually making them in order to be able to cope with its internal states and the external environment. The types involved in these judgements are intimately linked to agent's biology and embodiment as they are the types of basic representations generated by the sensors of the agent. As such, there is a finite set of these types. The assignment of other types is governed by the attention model of Load Theory. Attention can be either introduced by the task (or agent internally) or the context (agent externally). In terms of knowledge representation there is no difference between the types of the activity of tea making and the types associated with handling of dangerously hot objects. The set of task and context induced types for which an agent is primed at any moment

is defined by current pre-attentive judgements and the sequence of tasks and contexts the agent has been engaged so far. For example, given that the agent has previously been in the corridor coupled with new pre-attentive judgements could prime the agent to be attentive to types one typically judges in a kitchen. An agent learns through experience the types that are relevant in a particular task and context. Practically, this amounts to finding associations between types in agent's memory and their evolution over time.

3.2 Cognitive states

Thematic relations are relations between objects, events, people and other entities that co-occur or interact together in space and time (Lin and Murphy, 2001; Estes et al., 2011). Inspired by the concept of a thematic relation we propose that an agent's type inventory is organised as a set of cognitive states, where each state defines a set of types that are related by a thematic relationship. A cognitive state may be the cognitive correlate of the agent intentionally performing a task but may also be a non-explicit cognitive state of an agent generally being in a situation or having a disposition. Importantly, we don't believe that an agent has conscious access to all its cognitive states nor can all states be directly mapped to concrete activities. Rather a cognitive state can be understood as a sensitivity towards certain types of objects, events, and situations where this sensitivity mapping between states and entities has been learned from experience. For example, there may be a cognitive state associated with the agent's basic existence and its wish to continue existing, or of being a parent, or of being in a concert hall, or of being involved in a conversation about playing a trumpet, or of making a cup of tea. The commonality across this disparate set is the fact that it is possible to list a set of types that are relevant to each state which represents agent's resources in terms of (Cooper, 2008). For example, the very fact of an agent's existence makes it sensitive to entities in the environment that endanger it (large things moving towards it at speed) or help its existence (food nearby). The cognitive state related to being in a concert hall might prime the agent to make judgements about the music, the instruments and the conductor. The state of participating in a conversation about playing a trumpet prime judgements relating to the body language of your

interlocutor or the relationship you have with your interlocutor (are they an experienced player or an observer). Finally, the state of making tea could prime an agent to make judgements relating to kettles and cups and their arrangement in space.

It might appear that our approach simply pushes the intractability of judgements over the set of combinatorially exploding types onto the intractability over a set of cognitive states. We argue, however, that not withstanding the apparent complexity of human inner life there are in fact relatively restricted number of cognitive states that a human or an agent trying to live like a human needs to maintain in the course of an average day. While theoretically there could be as many states as the number of type judgements discussed earlier it is important to note that these states are built by an agent bottom up when an agent discovers new situations. Since an agent will be constrained by the environment in which it operates and since it can only discover a finite set of situations in its life, and since it is equipped with learning mechanisms with a strong bias to make generalisations it will only build a subset of these states that can be managed by its memory.

Important features of states and types include: (a) an agent may be in several states at the same time (they may be making tea and talking about music), and (b) a type may be associated with more then one state. While an agent is in a state or states performing any additional type judgements associated with one of the states incrementally reduces its ambiguity of being in several states.

4 A computational model

There are three requirements for our computational model: (i) agents clusters types according to thematic relations into several states, (ii) types are associated with each state with a certain probability, and (iii) a particular type may be associated with more than one state. Thematic relations between types are expressed by the co-occurrence of types in states. There are several computational mechanisms that could be used to automatically create states (clusters of types) with the above properties from data. For example, all three requirements are satisfied by *Latent Dirichlet Allocation (LDA)* which is a popular approach to topic modelling (Blei et al., 2003), the analogy between topic modelling and our scenario being that a topic is similar to a state and the association of a word

with a topic is equivalent to the association of a type with a state. A drawback with LDA is that the number of topics or states must be known a priori. However, *Hierarchical Dirichlet process* (Teh et al., 2006) is an extension of the LDA where the number of topics is also learned.

Given that an agent has learned thematic relations between types in the form of their association to states, the control problem which it is facing is that it cannot know which state it is in and consequently it cannot decide what is the optimal collection of types to be primed for in making judgements. As a result the agent must try to infer the best sets of types to prime for by estimating:

1. a posterior distribution over the possible states (and, updating this distribution as it receives observations from the world and makes judgements about the world)

2. make a decision regarding which judgements to be primed to make based on the updated probability distribution.

The posterior probability of being in a particular state at time t is dependent on the previous states at time $t-1$ (i.e., the Markov property holds: conditioned on the present the future is independent of the past), the task and context judgements the agent has made following the priming in the previous $t-n$ states where n is the length of history an agent keeps, and the new pre-attentive judgements which may reflect perceptual change in its world. So, the posterior probability of each of the cognitive states of the agent can be computed as follows:

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

where s_t is a state at time t, AS_{t-1} is the set of *active states*[1] at time $t-1$, Pre_t is the set of new pre-attentive judgements the agent has just made, $Task_{t-1}$ is the set of task relevant judgements the agent has made following previous priming, and *Cont* is the set of contextual judgements the agent has made following previous priming, and η denotes a normalisation process that ensures that the total probability mass of the posterior distribution sums to 1. We argue that the probabilities on the

---

[1]We will define the set of active states later.

right hand side of this equation can be learned from experience. This learning process can be simplified if we assume conditional independence between Pre_t, $Task_{t-1}$, $Cont_{t-1}$ and AS_{t-1} given s_t, essentially adopting the same formulation for calculating poster probabilities as is used by a standard naive Bayes' classifier:

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

Once we have computed a posterior probability over the set of states an agent has we need a mechanism that explains how this distribution informs the process of priming types. The simplest mechanism would be to select the state with the *maximum a posteriori* probability and then load into working memory the set of types that are associated with this state. This approach has the advantage of being computationally simple. However, it has the disadvantage that the agent assumes that they are only ever in one state, and, furthermore, if two or more states have a high posterior probability there is the possibility that the agent will keep switching between these states from one moment to the next. An alternative approach that is less susceptible to switching between states is to:

1. use the posterior probability over the states to rank and prune the state set, (the states that are not pruned are the *active states*)

2. renormalise the probability distribution over the set of active states,

3. compute a posterior probability over the set of types associated with active states using a *Bayes optimal classifier*,

4. using the posterior probability over types, rank and prune the set of types and load the set of unpruned types into working memory.

In order to rank and prune the state set we simply order the states based on their posterior probabilities and remove all the states that have a probability below a predefined threshold. This rank and pruning approach essentially implements a process whereby an agent can recognise what is not relevant to the current situation. Renormalising the probability distribution over the remaining

states is a simple process of summing the probability mass of the unpruned states and then dividing the probability mass of each state by this sum. The posterior probability of a type at time t is calculated using a Bayes optimal classifier as follows:

$$P(type_t|Pre_t,Task_{t-1},Cont_{t-1},AS_{t-1}) = \sum_{s\in AS_t} P(type|s) \times \quad (2)$$
$$P(s|Pre_t,Task_{t-1},Cont_{t-1},AS_{t-1})$$

where $type_t$ denotes a type at time t, AS_t denotes the set of unpruned (active) states at time t, $P(s|Pre,Task,Cont,AS_{t-1})$ denotes the probability of an active state s after the state set has been pruned and the posterior probability over the active states has been renormalised, and Pre_t, $Task_{t-1}$, $Cont_{t-1}$, and AS_{t-1} have the same meanings as defined above. Using a Bayes' optimal approach to calculating the posterior distribution over the types associated with the active states is computationally expensive because it includes a summation across the set of active states. However, the size of this set can be restricted based on the pruning criteria used so the computational cost of this summation operation can be minimised. Some of the benefits, however, of using a Bayes' optimal formulation are that: (a) this process explicitly recognises the fact that more then one state may be active at one point, (b) it also recognises the fact that a type may be associated with more then one state and that the strength of association between the type and a state is probabilistic ($P(type|s)$), and (c) this formulation is robust to small variations in the posterior distribution over states (i.e., when the state with the *maximum a posteriori* probability changes the system is stable—in terms of the types that are loaded into memory—if the changes across the distribution are stable). Once the posterior distribution over the types has been calculated the types can be ranked and pruned in a similar fashion to the states. This means that we need two thresholds for pruning, one for pruning the states and one for pruning the types. The ranking and pruning across the states and the types both reflect the attention based approach we have taken to this work modelled by Load Theory. When the cognitive load on the agent is low the pruning of states and types can be relaxed and when the cognitive load from the perceptual selection is high the pruning can become more severe.

5 Worked example

In this section we present a worked example that illustrates how an agent interacting in and moving around an environment can use the proposed models to prime the set of types judgements it has loaded in its memory. This example assumes that the agent has already learned a number of types and has already associated these types with the cognitive states it has constructed over the course of its lifetime.

To begin we will assume our agent has three cognitive states: $\mathcal{S}1$, $\mathcal{S}2$, $\mathcal{S}3$ and the prior probabilities of these states are $< 0.4, 0.3, 0.3 >$ respectively. Furthermore, the state transition matrix is a right stochastic matrix with i rows and j columns where each cell defines the probability of going from state i to state j in one time step (i.e., each cell defines $P(\mathcal{S}j|\mathcal{S}i)$) and is defined as follows:

	$\mathcal{S}1$	$\mathcal{S}2$	$\mathcal{S}3$
$\mathcal{S}1$	0.7	0.2	0.1
$\mathcal{S}2$	0.3	0.4	0.3
$\mathcal{S}3$	0.1	0.2	0.7

We will assume that there are 3 different pre-attentive types that the agent can assign to low-level perceptual features. For labelling convenience let us assume that these features are biased to particular locations in a building so that we can name these types after these locations, namely: OFFICE, CORRIDOR, and KITCHEN (cf. *semantic labelling of places* (Martínez Mozos et al., 2007)). We will also assume that the agent knows three task/contextual types[2]. We are interested in constructing agents that can participate in dialogues so we have decided that these types include types assigned to utterances in dialogue that the agent can engage in; for example, this agent can take part in dialogues relating to WEATHER, MACHINE-LEARNING, or the general WELL-BEING of someone. According to our model the agent should have learnt a probabilistic relationship between each of these types and its own cognitive states. The following right stochastic matrix defines the probabilistic relationship between each of the pre-attentive types and the state (i.e., each cell defines $P(type|\mathcal{S}i)$):

---

[2] For the purposes of the example the distinction between task and contextual types is moot.

	OFFICE	CORRIDOR	KITCHEN
$\mathcal{S}1$	0.7	0.2	0.1
$\mathcal{S}2$	0.1	0.8	0.1
$\mathcal{S}3$	0.05	0.15	0.8

And, the following matrix defines the probabilistic relationships between each of the task/context types and the states:

	WEATHER	MACH.-LEARN.	WELL-BEING
$\mathcal{S}1$	0.1	0.7	0.2
$\mathcal{S}2$	0.4	0.1	0.5
$\mathcal{S}3$	0.6	0.3	0.1

We also need to define two attention thresholds: one threshold is used to define the set of active states and the other is used to define the set of active types. Unlike the probabilities defined above (which are relatively fixed and are updated via a separate learning process) these attention thresholds may change from moment to moment and are dependent on the cognitive load the agent is experiencing: high load and the thresholds are low, low load and the threshold are high. For this example, we will assume that the agent is under a moderately high load and that both of these thresholds are set to: 0.3.

To begin calculating the set of types that the agent is primed for at time step t we need information relating to: (a) the set of active states at $t-1$ (AS_{t-1}); (b) the set of task and context type judgements the agent made at time $t-1$ (T_{t-1}); and, (c) the set of pre-attentive judgements the agent has just made at time t (Pre_t). For this example we will assume the following: $AS_{t-1} = \{\mathcal{S}1, \mathcal{S}2, \mathcal{S}3\}$, $T_{t-1} = \{\text{MACHINE-LEARNING}, \text{WEATHER}\}$, and $Pre_t = \{\text{OFFICE}, \text{CORRIDOR}\}$.

Our first step is to calculate the probability distribution over the states at time t. We do this using Equation 1. Before we apply Equation 1 we need to calculate the probability distribution for $P(AS_{t-1}|\mathcal{S}_t)$. We can calculate these probabilities using the prior probabilities for the states and the transition matrix $P(\mathcal{S}_j|\mathcal{S}_i)$ and applying Bayes' Theorem. The resulting probabilities (rounded to 4 places) are as follows:

	$\mathcal{AS}1_{t-1}$	$\mathcal{AS}2_{t-1}$	$\mathcal{AS}3_{t-1}$
$\mathcal{S}1_t$	0.7000	0.2250	0.0750
$\mathcal{S}2_t$	0.3077	0.4615	0.2308
$\mathcal{S}3_t$	0.1176	0.2647	0.6176

Once we have these probabilities it is a relatively straightforward process to calculate $P(\mathcal{S}_t|Pre_t, Task_{t-1}, Cont_{t-1}, AS_{t-1})$ using Equation 1. One technical point is that for each factor on the right hand-side of the equation ($P(Pre_t|s_t)$, $P(Task_{t-1}|s_t)$, $P(Cont_{t-1}|s_t)$ and $P(AS_{t-1}|s_t)$) we assume conditional independence between the conditioned events given the evidence. For example, for each $\mathcal{S}i_t$ we calculate $P(\mathcal{AS}_{t-1}|\mathcal{S}i_t)$ as:

$$
P(\mathcal{AS}_{t-1}|\mathcal{S}i_t) = P(\mathcal{AS}1_{t-1}|\mathcal{S}i_t) \\
\times P(\mathcal{AS}2_{t-1}|\mathcal{S}i_t) \times P(\mathcal{AS}3_{t-1}|\mathcal{S}i_t)
$$

In this context the posterior probability over the states (rounded to 4 places of decimal) is:

$$
\mathbf{P}(\mathcal{S}_t|Pre_t, Task_{t-1}, Cont_{t-1}, AS_{t-1}) = \\
< 0.5412, 0.3677, 0.0911 >
$$

Applying the attention threshold to this distribution over the states the set of active states at time t is then $\{\mathcal{S}1, \mathcal{S}2\}$ and renormalising the probability mass over these states gives us a probability distribution (again rounded to 4 places of decimal) of $< 0.5954, 0.4046 >$. We can now use Equation 2 to calculate the posterior probabilities over the task and contextual types. We only do this calculation for types that are associated (i.e., $P(type|\mathcal{AS}i) > 0$) with at least 1 of the active states. In this instance all three of the task/context types (WEATHER, MACHINE-LEARNING, and WELL-BEING) are associated with at least 1 of the active states so we calculate the posterior probability for all three of these types. The posterior probabilities over the types are $< 0.2214, 0.4573, 0.3214 >$. Applying the type attention threshold of 0.3 to this distribution there are two types that are active MACHINE-LEARNING, and WELL-BEING and the agent will be primed to make judgements of these types at time t.

In this example, we only pruned one of the task/context types from the primed list. However, as the number of types grows (remember that types will represent concepts at different levels of abstraction) and the number of states also grows then the number of types that are pruned will also grow.

6 Conclusion and future work

In this paper we present a computational mechanism for attention-driven type judgements in an

interacting agent that is inspired by cognitive processes in humans such as discovery of thematic relations and sharing of cognitive resources between perceptual selection and cognitive control as proposed in Load Theory. It is important to note that the problem of multiple type assignments or judgements is not exclusive to TTR but is a general problem where a cognitive agent has to make numerous classifications based on limited computational resources. In robotics this task is known as *visual search* (Sjöö, 2011; Kunze et al., 2014). The proposed application of TTR allows us to formulate a cognitively-inspired computational model for visual search. The approach is also relevant to computational modelling of situated dialogue. Being primed for particular types would disambiguate interlocutors utterances based on the previous type judgements and perceptual observations. In dialogue generation it allows priming of the agent to particular topics and therefore can be used for topic modelling of a dialogue system.

The model proposes that an agent has a set of cognitive states that they have learned from past experience. An agent may be in more than one cognitive state at any one time. There are a set of types associated with each cognitive state of an agent. When a cognitive state is active (unpruned) an agent is primed to make judgements relating to the types associated with the state. This is why our account links judgement in TTR and attention. The difficulty with this account is that because more than one cognitive state is active at any one time the agent must decide which of the active cognitive states it should prime for observation. The solution is that the agent should maintain a distribution over its cognitive states and prime its observation relative to the types associated with the cognitive states with high probability. Following Load Theory the agent will actually perceive as many of these primed types as it can before its perceptual capacity is exhausted and it will then select a subset of these primed types for further cognitive processing.

In our forthcoming work we are working towards a computational application of the model to situated dialogue. We are particularly interested in evaluating the benefits of an agent being primed this way in comparison to when it has no priming at all. The model introduces several parameters, for example the number of states, the number of types, the size of memory for pre-attentive judgements and task and context related judgements whose effects on system performance will also be investigated.

References

D. Alan Allport, Barbara Antonis, and Patricia Reynolds. 1972. On the division of attention: A disproof of the single channel hypothesis. *The Quarterly journal of experimental psychology*, 24(2):225–235.

Patrick Blackburn and Johan Bos. 2005. *Representation and inference for natural language. A first course in computational semantics*. CSLI Publications.

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

D.E. Broadbent. 1958. *Perception and Communication*. Pergamon Press.

Herbert H. Clark. 1996. *Using language*. Cambridge University Press, Cambridge.

Robin Cooper, Simon Dobnik, Shalom Lappin, and Staffan Larsson. 2015. Probabilistic type theory and natural language semantics. *Linguistic Issues in Language Technology — LiLT*, 10(4):1–43, November.

Robin Cooper. 2005. Austinian truth, attitudes and type theory. *Research on Language and Computation*, 3(2):333–362.

Robin Cooper. 2008. Type theory with records and unification-based grammar. In Fritz Hamm and Stephan Kepser, editors, *Logics for Linguistic Structures. Festschrift for Uwe Mönnich*, volume 201, page 9. Walter de Gruyter.

Robin Cooper. 2012. Type theory and semantics in flux. In Ruth Kempson, Nicholas Asher, and Tim Fernando, editors, *Handbook of the Philosophy of Science*, volume 14 of *General editors: Dov M Gabbay, Paul Thagard and John Woods*. Elsevier BV.

M. W. M. G Dissanayake, 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.

David R Dowty, Robert Eugene Wall, and Stanley Peters. 1981. *Introduction to Montague semantics*. D. Reidel Pub. Co., Dordrecht, Holland.

John Duncan. 1980. The locus of interference in the perception of simultaneous stimuli. *Psychological review*, 87(3):272–300.

Zachary Estes, Sabrina Golonka, and Lara L Jones. 2011. Thematic thinking: The apprehension and consequences of thematic relations. In Brian Ross, editor, *The Psychology of Learning and Motivation*, volume 54, pages 249–294. Burlington: Academic Press.

Jonathan Ginzburg and Raquel Fernández. 2010. Computational models of dialogue. In Alexander Clark, Chris Fox, and Shalom Lappin, editors, *The handbook of computational linguistics and natural language processing*, Blackwell handbooks in linguistics, pages 429–481. Wiley-Blackwell, Chichester, United Kingdom.

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

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

Daniel Kahneman. 1973. *Attention and effort*. Prentice-Hall, Englewood Cliffs, N.J.

Geert-Jan M. Kruijff, John D. Kelleher, and Nick Hawes. 2006. Information fusion for visual reference resolution in dynamic situated dialogue. In *Perception and Interactive Technologies*. Springer.

Lars Kunze, Chris Burbridge, and Nick Hawes. 2014. 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, March, 24–26.

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

Nilli Lavie, Aleksandra Hirst, Jan W de Fockert, and Essi Viding. 2004. Load theory of selective attention and cognitive control. *Journal of Experimental Psychology: General*, 133(3):339–354.

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

Óscar Martínez Mozos, Rudolph Triebel, Patric Jensfelt, Axel Rottmann, and Wolfram Burgard. 2007. Supervised semantic labeling of places using information extracted from sensor data. *Robotics and Autonomous Systems*, 55(5):391–402, 5.

Kristoffer Sjöö. 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.

Luc Steels and Tony Belpaeme. 2005. Coordinating perceptually grounded categories through language: A case study for colour. *Behavioral and Brain Sciences*, 28(4):469–489.

Yee Whye Teh, Michael I. Jordan, Matthew J. Beal, and David M. Blei. 2006. Hierarchical dirichlet processes. *Journal of the American Statistical Association*, 101(476):1566–1581.

Anne M. Treisman. 1969. Strategies and models of selective attention. *Psychological review*, 76(3):282–299.

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

A.T. Welford. 1967. Single-channel operation in the brain. *Acta Psychologica*, 27:5–22.