The Integrated Model of Memory: a dialogical perspective

Jonathan Ginzburg[1] **and Andy Lücking**[1,2]
[1]Université Paris Cité, CNRS,
Laboratoire de Linguistique Formelle (UMR 7110)
[2]Goethe University Frankfurt
yonatan.ginzburg@u-paris.fr, luecking@em.uni-frankfurt.de

Abstract

The increasing complexity of dialogue information states raises the question of their ontological status. To this foundational question one can add a more concrete concern: all existing semantic frameworks for dialogue while designed to explain how meaning emerges from the 'accumulation of information', have no corresponding means of *eliminating* information. Our claim, which we exemplify, is that *memory boundedness impacts dialogue coherence*. This paper aims to offer an initial sketch of an approach that both resolves the foundational issue raised above and the issue of memory fragility. We propose to construe dialogue information states as properties of brain networks. This follows in the programme of brain-grounded semantics (Hagoort, 2020). Our strategy involves taking a recent framework for describing the dynamics of memory (Bastin et al., 2019) as a basis for developing a suitable notion of cognitive states and their dynamics for dialogue interaction. We sketch a semantic description of this system, suggesting that this imposes strict conditions on potential semantic frameworks.

1 Introduction

All contemporary semantics for dialogue are *dynamic*: they view many aspects of meaning as emerging from context change. But whereas 'context' was an inert, abstract notion in early Montague semantics (Montague, 1974) and an eventuality in situation semantics (Barwise and Perry, 1983), dynamic semantics starting with Discourse Representation Theory (DRT) (Kamp, 1981) identified contexts with *information states*. Whereas originally such information states tracked discourse referents and presuppositions, in recent work on dialogue information states have become complex as a wide range of phenomena have been analyzed, including the visual field (Lücking, 2016) (for analyzing manual gesture), emotional structure (Ginzburg et al., 2020) (for analyzing laughter), and defeasible common sense knowledge (*topoi/enthymemes* (Breitholtz, 2020) (for analyzing rhetorical relations). While there seems little doubt that this range of information is used in dialogue interaction, it does raise the question what kind of entity encompasses all these diverse types of information. What *is* the *dialogue gameboard* (DGB) posited in frameworks like KoS (Ginzburg, 2012)?

One is free to adopt a Cartesian perspective, as has often been the case in Chomskyan theoretical linguistics, though this is arguably an avenue that leads to untestable modelling (Poeppel and Embick, 2005). To this foundational question one can add a more concrete concern: all existing semantic frameworks for dialogue while designed to explain how meaning emerges from the 'accumulation of information', have no corresponding means of *eliminating* information—there are operations in DRT that make discourse referents inaccessible and KoS has notions of downdating questions, but long-term information established as accepted, is locked in for ever more. This means that, as Ginzburg and Lücking (2020) put it, 'forgetting is forgotten'—there is no natural way to deal with the fragility of memory, an intrinsic and concrete feature of human interaction, both involving neurotypicals and non-neurotypicals like dementia sufferers. Our claim, exemplified below in section 2, is that *memory boundedness impacts dialogue coherence*.

This paper aims to offer an initial sketch of an approach that both resolves the foundational issue raised above and the issue of memory fragility. The basic idea is straightforward, namely to construe dialogue information states as properties of brain networks (Bressler and Menon, 2010). This follows in the programme of brain-grounded semantics (Hagoort, 2020). This emphasizes the need to ground semantics in brain–internal processes, while ensuring that top-down causation (coming from the computational level, in this case, say the DGB) is given its due (Campbell, 1974). Thus, in Marrian

Proceedings of the 26th Workshop on the Semantics and Pragmatics of Dialogue, August, 22-24, 2022, Dublin.

terms, this does not mean in any way downgrading the computational level of explanation, as provided by semantic theories of dialogue, but ensuring that this is commensurate with the algorithmic (and ultimately) implementational levels.[1]

Our strategy will be to take a recent framework for describing the dynamics of memory (Bastin et al., 2019), which we survey in section 3, as a basis for developing a suitable notion of cognitive state for dialogue interaction. In section 4 we sketch a semantic description of this system, suggesting that this imposes strict conditions on potential semantic frameworks—requiring probabilistic judgements and operations adding and removing structure from representations; in contrast to the passive view implicit in the lab encloistered memory literature, recollection processes are constituents of interactions, giving rise to clarification interaction, laughter, and crying. We exemplify the framework with reference to our earlier examples in section 5. This does not mean a behaviorist account eschewing unobservables, but an attempt to formulate theory in a way that is *ceteris paribus* consistent with current observations about brain geography and dynamics.

We build on an earlier work (Ginzburg and Lücking, 2020) that tried to forge a link between dialogue semantics and theories of memory. In particular, the assumption that DGBs are constituents of *episodic memory*. The emphasis in the earlier paper was on short-term memory aspects of dialogue, which are indeed the most salient aspects needed for dialogue processing (resolution of indexicals, non-sentential utterances, disfluencies etc), though the paper also addressed long-term aspects. We will concentrate on the latter here while offering some significant modifications to the earlier account. That account was primarily a formalization of a Baddeley style architecture (Baddeley, 2012), which is highly motivated empirically, but has no pretensions to direct brain realization (Hasson et al., 2015). We will not assume a dichotomous short/long-term distinction, but follow, e.g., Hasson et al. (2015) by assuming that such differences can be captured in terms of short/long temporal receptive windows (Kiebel et al., 2008; Golesorkhi et al., 2021), a view which is also consistent with recent work that suggests that time-dependent forgetting across both short and long terms is related to degradation of hippocampal-dependent relational information (Sadeh and Pertzov, 2020).

2 Memory and Dialogue Coherence: some data

Consider first (1). The initial laughs by A and B, as suggested by Ginzburg and Lücking (2020), arise as a consequence of the clash between the observed visual scene and the topos *presidents wear formal suits*. Now consider B's second laugh a year later: this is ambiguous between a laugh about the incongruity of the recollected event of viewing Putin or a pleasure laugh about the autobiographical event a year before. This can only be explained by appeal to episodic memory (and semantic memory for the topos), distinctions unavailable in standard dynamic semantic treatments of context.

(1) A and B observe Putin wearing a hazmat suit on tv:[2]

A: laughs
B: laughs
[*A year later:*]
A: Do you remember that bizarre situation with Putin during Covid?
B: laughs

(2) is an apocryphal story about the mathematician Paul Erdös. This illustrates a basic feature of conversational interaction, namely that this involves an initial check whether the interlocutor is familiar or not; familiarity requires an initial intimacy interaction, whereas lack of familiarity (as here) an establishment of the interlocutor's identity:

[1] We hope this provides at least a partial answer to a worry expressed by an anonymous SemDial reviewer 'I don't see why we can't leave the modeling of when something is accessible through memory, and for how long, to the cognitive scientists and then on the linguistic side pick up the ball once it has been determined that there is or is not a referent.'

[2] Kremlin.ru, CC BY 4.0, https://commons.wikimedia.org/wiki/File:Vladimir_Putin_in_Kommunarka_hospital1.jpg.

(2) ERDÖS: Where are you from?
MATHEMATICIAN: Vancouver
ERDÖS: Really? Then you must know my friend Elliot Mendelson.
MATHEMATICIAN: (pause) I am your friend Elliot Mendelson.

(3) illustrates forgetting on a short time scale so requires a means of explaining short-term lack of recollection for an event, along with the attendant potential for repair:

(3) CAROL: Suddenly this means a lot to them. Yes? / Critical illness cover, that's great. Excuse me a minute. (Knocking at the door)
UNKNOWN1: Sorry to interrupt, I've come to collect the packet. /
CAROL: Oh right, it's the bag, sorry there isn't one tonight. /
UNKNOWN1: See you then /
CAROL: Thanks for coming then, yes, bye. That's good, I forgot the post. **Erm, where was I? What was I talking about?** /
UNKNOWN2: Single people. (BNC)

(4) and (5) both involve dementia sufferers (participants (PAR) interacting with an investigator (INV)).[3] In (4) there is explicit reference to a failed word recollection; in (5) the speaker makes reference to a descriptively potent but name–lacking cognitive state. In both cases the speakers' cognitive states maintain social norms relating to embarrassment which underwrite the laughter, but (5) in particular exhibits depression, characteristic of dementia sufferers, in part caused by repeated memory failure.[4]

(4) Becker et al. (1994), Pitt corpus, fluency 043-0, 04–10
INV: I want you to tell me as many animals as you can think of in one minute „ okay?
INV: they can be farm animals or zoo animals or pets.
INV: they can't be birds or fish or insects „ okay?
INV: can you begin?
PAR: &=laughs no. [+ exc]
INV: +< no?
PAR: +< (be)cause I forgot. [+ exc]

(5) DePaul (2017), depaul2a, 12–15
PAR: I can picture &=points:forehead whatever things that I'm still seeing or whatever.
PAR: but I don't know what to call it.
PAR: that's [/] that's what's whatever.
PAR: when I go to heaven it's gonna be &=looks:down &=head:shakes fine &=laughs.

(6) illustrates that a successfully recalled event involves reappraisal:

(6) (Interview with Pete Doherty about his relationship with Barât—they had one of the most fractious relationships in rock music. . .)
JOURNALIST: And yet the intensity of your bond was palpable.
DOHERTY: Absolutely. You're making me quite emotional my eyes are filling with tears. (The Guardian, June 2022).

3 The Integrated Model of Memory

In this section we summarize the Integrative Model of Memory (IMM) (Bastin et al., 2019), a synthetic effort to incorporate recent neuropsychological models of memory. We use this framework as a basic description of relevant brain networks.

3.1 Basic phenomena

The two main phenomena the theory attempts to explain are the brain processes which give rise to (event) *recollection* and (entity) *familiarity*, examples of which we saw in section 2.

[3] Transcription follows the CHAT format (MacWhinney, 2000). The symbol "[/]" indicates a repeated attempt to produce a word, double comma "„" is an interactional marker for an intonational group, the symbol "&=" prefixes speaker actions, "+<" indicates a slight overlap of utterances, parentheses enframe omitted material, "[+ exc]" is a user-defined postcode which marks utterances that are excluded from analyses.

[4] Although the relationships between dementia and depression are complex, e.g., Bennett and Thomas, 2014.

3.2 Basic explanatory mechanisms

IMM relies on a combination of distinct types of representations and processes in its explanation of recollection and familiarity. As far as representations go, it distinguishes between the following kinds of representations:

- Event representations: representations of relations between two or more entities. These representations are associated with the hippocampus. Following indexing theory (Teyler and Rudy, 2007), such representations do not actually store a detailed representation of the event, but an index that enables to retrieve from the neocortex the original modes in which the event was perceived (visual, aural etc).

- Entity representations. These are representations often arising in a one shot manner (Kent et al., 2016) in the perirhinal/anterolateral entorhinal cortex that allow the discrimination of objects with overlapping features such as faces in a viewpoint-invariant manner (Erez et al., 2016). This enables quick recognition of familiar objects in the stream of objects perceived in the environment. This location is also one where conceptual features may get bound to entities via interaction with the anterior temporal area (Martin et al., 2018). The entity-level representations in the anterolateral entorhinal/perirhinal cortex correspond to a higher level of representation of the object, representing the individual object in a way abstracted from its presentation characteristics (viewpoint, perceptual conditions of presentation etc.)

- Background representations:[5] a network linking the parahippocampal/posteromedial entorhinal cortex and the occipitoparietal cortex and retrosplenial cortex. This system provides the setting into which entities fit in within events, binding the two enables entities to gain distinct significance based on diverse settings.

Key processes in the IMM are:

- Pattern separation (Rolls, 2016; Ngo et al., 2021): a hipppocampal process in which similar inputs are given separate representations based on specific conjunctions of features.

- Pattern completion (Rolls, 2016; Ngo et al., 2021): a hipppocampal process by means of which a partial information cue triggers the reactivation of the complete pattern.

- Attribution mechanisms (Whittlesea and Williams, 2000): recollection and familiarity are not merely determined by the accuracy of representations, but by task-dependent confidence thresholds. In the highest band are commonly encountered entities whose familiarity is automatic, in the lowest band unknowns; the middle band consists of entities whose recognition triggers incongruity—this incongruity is the subjective feeling of fluency. Seeing a person resembling a work colleague will lead to different judgements and actions depending on whether I need to decide if to greet him or merely to report seeing him.

The representational structure of the IMM is summarized in Fig. 1.

3.3 Unofficial Extensions: Semantic Memory and Emotion

The IMM is an ambitious programme, but in its initial formulation at least (Bastin et al., 2019), it makes some understandable simplifying assumptions. We mention here two, which we think need to be eliminated for the viability of a linguistically oriented theory, using suggestions in Bastin et al. (2019) and in responses to the paper.

The IMM considers only episodic memory. But as argued in Greenberg and Verfaellie (2010) and, building on this, by Gainotti (2019), there is an intrinsic dependence between this system and what has been called *semantic memory*—"the memory necessary for the use of language" (Tulving, 1972, p. 386). There is ample evidence of disassociation between the two—medio temporal lobe (MTL) damage can severely hinder the subsequent formation of episodic memories without affecting semantic memory (Scoville and Milner, 1957), whereas semantic dementia, which leads to loss of naming ability, can have minor effects on episodic memories (Chan et al., 2001). Nonetheless, there is evidence that semantic memory facilitates the acquisition of new episodic memories and vice

[5] In the IMM these are called 'context representations' given that much of the experimental data derives from manipulation of cards with visual images. Since 'context' plays a major role in formal semantics, we have changed the terminology.

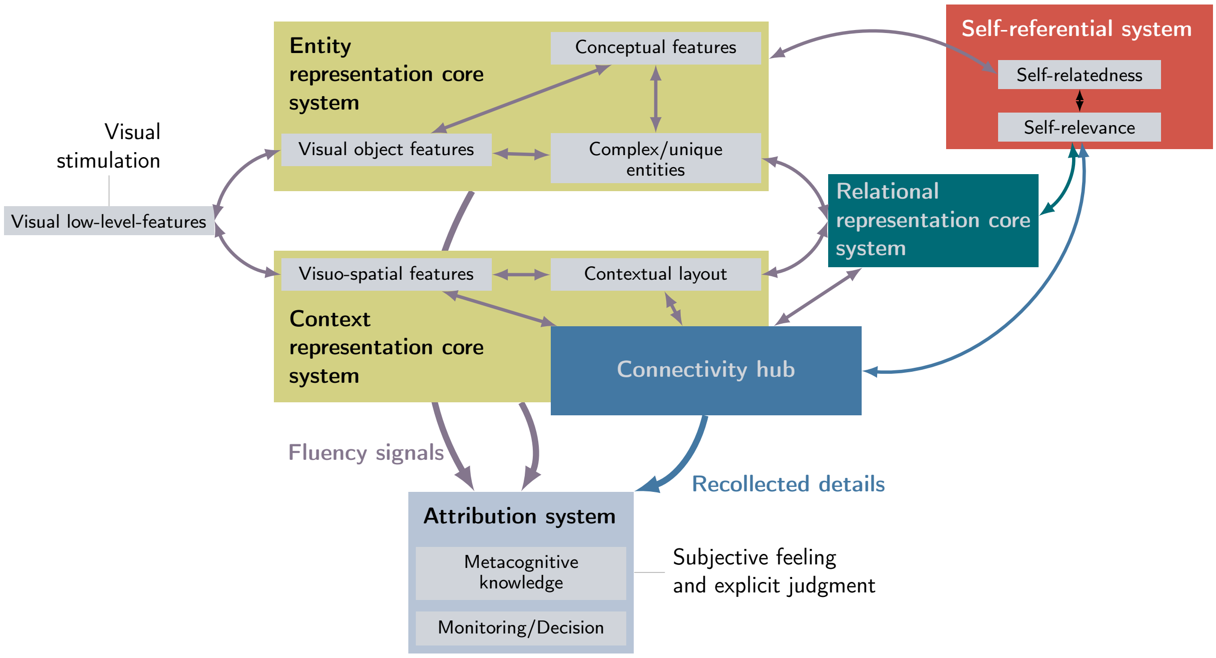

Figure 1: Slightly shortened illustration of the IMM, adopted from Bastin et al. (2019, p. 4).

versa (Greenberg and Verfaellie, 2010). Conversely, episodic memories facilitate the retrieval of information from semantic memory, and semantic memories constitute an important base from which complex and detailed episodic memories are constructed. The distinction between episodic and semantic memory is not straightforward and goes against a dichotomous explication (somewhat reminiscent of the stage level/individual level distinction in semantics (Carlson, 1977)). Tulving (1972) suggested a serial encoding hypothesis (perceptual → semantic → episodic), but the finding that episodic memory can facilitate new semantic learning is harder to reconcile with this, while it is consistent with the view of semantic memory as decontextualized episodic memory (Baddeley, 1988). However, as Greenberg and Verfaellie (2010) argue, the fact that degeneration of semantic memory is correlated with a severely weakened and vague episodic memory does not cohere well with a notion of parallel storage; a more attractive view is that episodic memory effectuates a binding between contextual information and material found in semantic memory.

The second simplification inherent in the current version of the IMM concerns its abstracting away from emotion. Already in the 1970s there was evidence that positively valenced events are remembered at a higher rate (Kintsch and Bates, 1977); there is much more recent evidence that emotional memories are forgotten at a slower rate than neutral memories over long timescales (e.g., a day vs. 5 min; Sharot and Yonelinas, 2008). Yonelinas and Ritchey (2015) argue that the slower forgetting of emotional memories can be linked to a dependence on the amygdala and its interaction with nonhippocampal MTL structures, rather than on the hippocampus. This hypothesis aligns well with the notion, promoted in Sadeh and Pertzov (2020), that hippocampal representations are more prone to temporal degradation than nonhippocampal representations.

3.4 Applications to memory deterioration

3.4.1 Neurotypical forgetting

One account of forgetting links it to *contextual drift* (Yonelinas et al., 2019). On this view forgetting as evinced in lab settings arises from a change or drift in context between study and test. Furthermore, on this approach, forgetting may be further promoted by *contextual interference*, such as intervening activities or physical changes. Manohar et al. (2019) suggest that memory encoding depends on rapid plasticity in flexibly coding neurons that may reside in the hippocampus. Such plasticity allows distinct representations that give rise to binding which results in a coherent memory representation. Time-dependent volatility of the synaptic weights is expected to lead to forgetting of relational or conjunctive information over time. Such forgetting does not occur "because of any specific decay rule, but rather because the plasticity rule operates continuously to alter all synaptic weights, and this 'erodes' the representations that are not currently active" (Manohar et al., 2019).

On the neuronal level, stored activation patterns (i.e. memory) are subject to three kinds of persistence-affecting processes, namely (i) neurogenesis (that is the creation of new neurons in, e.g., the hippocampus), (ii) synaptic weight decay, and (iii) synapse elimination (Richards and Frankland, 2017, p. 1072). As a consequence, memory even of neurotypical beings is a "transient" affair.

3.4.2 Neuroatypical memory failure

Neuroatypical characteristics may reinforce the afore-mentioned neuronal processes of synaptic (in-)stability. According to the IMM, the dissociation of recollection and familiarity in patients with lesions selective to the hippocampus or perirhinal/entorhinal cortex (e.g., Barbeau et al. (2011)) arises because the core representations are damaged. Clinical evidence validating these predictions is discussed in Bastin et al. (2019), in particular with respect to Alzheimer Disease.

4 Integrating Dialogue Semantics and Memory

In this section we introduce basic notions of KoS, which exemplifies a theory of dialogue states and their dynamics (at a computational level). We then sketch how this theory can be construed in terms of memory structures (at an 'algorithmic level').

4.1 Dialogical Cognitive States

KoS (Ginzburg, 2012; Ginzburg et al., 2020)—formulated using the logical framework TTR (Cooper and Ginzburg, 2015; Cooper, 2022)—is a theory of dialogue that offers an account of how speech events and other multimodal meaning bearing events change an individual's cognitive state. Instead of assuming a single context to be operative, a collective notion is emergent from individual *Total Cognitive States* (TCS), one per participant. A TCS has two partitions, namely a *private*—about which we will not elaborate here—for details see (Larsson, 2002), and a *public* one, the DGB.

(7) $$\text{TCS} =_{def} \begin{bmatrix} \text{public} : \textit{DGBType} \\ \text{private} : \textit{Private} \end{bmatrix}$$

Dialogue gameboards (see 8a for the basic structure) track various aspects of the emerging context The parameters *spkr* and *addr* together with the addressing condition (at a given time) track verbal turns and mutual engagement; *Vis-sit* represents the visual situation of an agent, including his or her focus of attention (*foa*), which can be an object (*Ind*), or a situation or event (*Sit*), relevant *inter alia* for processing gestural answers; *facts* represents the shared assumptions of the interlocutors; uncertainty about mutual understanding that remain to be resolved across participants—*questions under discussion*—are a key notion in explaining coherence and various anaphoric processes (Ginzburg, 2012; Roberts, 1996) and is tracked by the parameter *qud*; dialogue moves that are in the process of being grounded or under clarification are the elements of the *pending* list; already grounded moves are moved to the *moves* list; finally, *mood* represents the publicly accessible emotional aspect of an agent that arises by publicly visible actions (such as non-verbal social signals, as well as by verbal exclamations), which can but need not diverge from the private emotional state; the result of appraisals is given in terms of structures like (8b) (Russell, 2003).

(8)

a. $$\textit{DGBType} =_{def} \begin{bmatrix} \text{spkr} & : \textit{Ind} \\ \text{addr} & : \textit{Ind} \\ \text{utt-time} & : \textit{Time} \\ \text{c-utt} & : \text{addressing(spkr,addr,utt-time)} \\ \text{facts} & : \textit{Set(Prop)} \\ \text{vis-sit} & = \begin{bmatrix} \text{foa} : \textit{Ind} \vee \textit{Rec} \end{bmatrix} : \textit{RecType} \\ \text{pending} & : \textit{List(LocProp)} \\ \text{moves} & : \textit{List(IllocProp)} \\ \text{qud} & : \textit{poset(Question)} \\ \text{mood} & : \textit{Appraisal} \end{bmatrix}$$

b. $$\textit{Appraisal} =_{def} \begin{bmatrix} \text{pleasant} : \begin{bmatrix} \text{Pred} = \textit{Pleasant} : \textit{EmotivePred} \\ \text{affect} : \begin{bmatrix} \text{pve} : \mathbb{N} \\ \text{nve} : \mathbb{N} \end{bmatrix} \end{bmatrix} \\ \text{responsible} : \textit{RecType} \\ \text{power} : \begin{bmatrix} \text{Pred} = \textit{Powerful} : \textit{EmotivePred} \\ \text{control} : \mathbb{N} \end{bmatrix} \end{bmatrix}$$

Conversational rules are the means for specifying how DGBs evolve. The types specifying its domain and its range we dub, respectively, the *pre(conditions)* and the *effects*, both of which are subtypes of DGBType: they apply to a subclass of records that constitute possible DGBs and modify them to records that constitute possible DGBs. Conversational rules are written here in a form where the preconditions represent information specific to the preconditions of this particular interaction type and the effects represent those aspects of the preconditions that have changed.

KoS can represent locutionary, (9a,b), illocutionary updates, as in (9c,d), and emotion-based updates, such as (9e):

(9) a. Utterance integration: an utterance is perceived, updates Pending as a *locutionary proposition* (a record consisting of a representation of the utterance u and a grammatical type T_u calculated to classify it); there is then an attempted instantiation of the contextual parameters of T_u; if successful, the locutionary proposition is updated with the contextual instantiation and an attempt is made to find an appropriate Move update rule; if successful, Moves gets updated; otherwise repair ensues: the utterance remains in Pending and a clarification question is calculated and posed.

Clarification question: if A's utterance u is in Pending, QUD can be updated with the question *What did A mean by u*.

b. Ask/Assert QUD-incrementation: given a question q and ASK(A,B,q)/Assert(A,B,p) being the LatestMove, one can update QUD with q/p? as MaxQUD.

c. QSPEC: this rule characterizes the contextual background of reactive queries and assertions—if q is MaxQUD, then subsequent to this either conversational participant may make a move constrained to be q-specific (i.e., either a direct answer or a sub–question of q).

d. Positive affect incrementation of Mood: given the LatestMove being an incongruity proposition by the speaker, the speaker increments the positive pleasantness recorded in Mood to an extent determined by the laughter's arousal value.

The latter rule, which will play some role below, can be formalized as in (10)—updates are weighted between new and old values using the weight ε:[6]

(10) PositivePleasantnessIncr$(\delta,\varepsilon) =_{def}$

$$\left[\begin{array}{ll}\text{preconditions:} & \left[\text{LatestMove.cont} : \textit{IllocProp}\right] \\ \text{effect :} & \left[\begin{array}{l}\text{Mood.pleasant.affect.pve} = \\ \varepsilon(\text{preconds.Mood.pleasant.affect.pve}) + \\ \quad (1-\varepsilon)\delta : \textit{Real} \\ \text{Mood.pleasant.affect.nve} = \\ \varepsilon(\text{preconds.Mood.pleasant.affect.nve}) : \\ \quad \textit{Real}\end{array}\right]\end{array}\right]$$

4.2 Dialogical Cognitive States and Memory Dynamics

Our starting point towards integrating dialogical cognitive states in memory is the idea from Ginzburg and Lücking (2020) that conversations are elements of episodic memory, which for concreteness we will assume are structured by DGBs. Whereas Ginzburg and Lücking (2020) considered short-term memory, within a Baddeley-style WM approach, we will not consider such aspects here, hence short-term elements relating to perception such as *Pending* (corresponding to the phonological loop) and *VisualSituation* (corresponding to the visuo-spatial sketchpad) are not included. What remain is specified by the type L(ong-term)DGBType, given in (11a).[7] Hence, we assume episodic memory track such episodes, as in (11b):

(11)

a. LDGBType $=_{def}$

$$\left[\begin{array}{l}\text{participants} = \left\{\text{x,y}\right\}: \textit{Set(Ind)} \\ \text{Moves : } \textit{List(LocProp)} \\ \text{QUD : } \textit{Poset(Question)} \\ \text{Mood : } \textit{Appraisal}\end{array}\right]$$

b. Episodic $=_{def}$

$$\left[\text{Conversational : list(LDGBType)}\right]$$

We distinguish several distinct types of memory representations. Events are perceived visually or aurally or often multimodally. We assume such

[6] *NegativePleasantnessIncr* is the analogous operation incrementing the .nve and .pve values of pleasantness *mutatis mutandis*.

[7] Eliminating *Pending* and *VisualSituation* from LDGBType is a simplifying assumption. There clearly has to be some representation of the perceptible visual scene during a conversation as part of its recollection. This issue relates to the fundamental issue of how short-term memory structure relates/maps onto long-term memory structure which we plan to address in an expanded version of this paper.

events are represented by structured, relational representations—formally via TTR record types (Cooper, 2022); the tokens are the external, real world manifestations of the internal types.[8] Events undergo appraisal which leads to both updates in the current emotional makeup of the cognitive state (see the type *Appraisal* above) and to creating episodic indices in the hippocampus, which are in effect vertices in a network connecting to percepts of events stored neocortically. We assume that such indices are created for events with positive pleasantness above a threshold or negative pleasantness above a larger threshold—which yields a bias for long-term memory of enjoyable events or of highly unpleasant ones. The rule in (12) creates a fresh index and associates it with the current pending event ("HC" abbreviates *hippocampus*):[9]

(12) HC index creation

$$\begin{bmatrix} \text{preconds}: & \begin{bmatrix} \text{Pending} : \textit{RecType} \\ \text{c1 : Pending.Mood.pleasant.affect.pve} \\ \geq \theta_1 \\ \quad \vee \text{ Pending.Mood.pleasant.affect.nve} \\ \geq \theta_2 \end{bmatrix} \\ \text{effects}: & \begin{bmatrix} \text{n = card(HC-Indices)+1}: \mathbb{N} \\ \text{HC-indices} := \text{HC-Indices} \cup \left\langle \text{n,Pending} \right\rangle \end{bmatrix} \end{bmatrix}$$

Both entity and semantic memory representations[10] are modelled as record types whose external witnesses correspond to real world individuals and (spatio-temporally unlocated) facts about these. We assume these arise from event percept (representations) by record type projection. We do not offer here general definitions, merely exemplify for the entity case:

(13) Entity representation creation:

a. Input:
$$\begin{bmatrix} \text{x} : \textit{Ind} \\ \text{C} : \textit{faceshape} \\ \text{c1 : C(x)} \\ \text{c}_{name} : \text{Name(Emmo,x)} \\ \text{y} : \textit{Ind} \\ \text{c2 : Hammer(y)} \\ \text{t} : \textit{Time} \\ \text{c3 : Hold(x,y,t)} \end{bmatrix}$$

b. Output:
$$\begin{bmatrix} \text{x} : \textit{Ind} \\ \text{C}: \textit{faceshape} \\ \text{c1 : C(x)} \\ \text{c}_{name} : \text{Name(Emmo,x)} \end{bmatrix}$$

Building on the discussion in section 3, we can describe the process for testing whether an entity is familiar. For simplicity we assume that the parameter used by the attribution system is relativized by the maximal element of QUD, though clearly this is a more intricate, domain sensitive (range of) parameter(s):[11]

1. Given an entity of type T_{source}, one searches in *Entities* for a match, a type T_{target} such that $T_{source} \sqsubset T_{target}$.

2. If one finds T_{target} such that $prob(match(T_{source}, T_{target}, \text{MaxQUD})) \geq \theta_{high}$, then $known(T_{source}.x)$.

3. If one finds T_{target} such that $\theta_{high} \geq prob(match(T_{source}, T_{target}, \text{MaxQUD})) \geq \theta_{low}$, then $familiar(T_{source}.x)$.

4. If all potential matches are evaluated as $\theta_{low} \geq prob(match(T_{source}, T_{target}, \text{MaxQUD}))$, then $\neg familiar(T_{source}.x)$

Given this notion of familiarity, we can sketch the process of *familiarity testing* that occurs as an interaction is initiated, resulting either in the latest-move (l-m) being an initial pleasantry or identity clarification:[12]

(14) Familiarity witnessing

a.
$$\begin{bmatrix} \text{preconds}: & \begin{bmatrix} \text{moves} = \left\langle \right\rangle \\ \text{addr} : \textit{Ind} \\ \text{c1 : familiar(addr)} \end{bmatrix} \\ \text{effects}: & \begin{bmatrix} \text{l-m.cont} : \textit{IllocProp} \\ \text{q} : \textit{Question} \\ \text{c2} : \textit{Recent-common-experience}\text{(q)} \\ \text{c3}: \textit{Co-Propositional}\text{(l-m.content,q)} \end{bmatrix} \end{bmatrix}$$

[8]Though of course misperception/delusion can lead to representations without external counterparts.

[9]Although we do not spell this out here, we could postulate additional binding to the amygdala in case of strong emotional arousal, both negative and positive (Maren and Quirk, 2004; Phelps, 2004). This would capture the fact that it is much more difficult to forget highly emotional events since the amygdala is more stable than the HC.

[10]It is not impossible to have episodic metalinguistic memories, but not the norm.

[11]We thank an anonymous SemDial reviewer for a subtle but important correction of stage 2 of the process.

[12]Here two utterances are CoPropositional if the questions (construed as propositional functions) they update QUD with (see rule 9b) have overlapping ranges (answers); for instance 'Whether Bo left', 'Who left', and 'Which student left' (assuming Bo is a student.) are all co-propositional.

b. $$\left[\begin{array}{l} \text{preconds} : \left[\begin{array}{l} \text{moves} = \langle\rangle \\ \text{addr} : \textit{Ind} \\ \text{c1} : \neg\text{familiar(addr)} \end{array}\right] \\ \text{effects} : \left[\begin{array}{l} \text{l-m.cont} : \textit{IllocProp} \\ \text{q} = \text{?Identity(addr)} : \textit{Question} \\ \text{c3}: \textit{Co-Propositional}(\text{l-m.content,q}) \end{array}\right] \end{array}\right]$$

Finally, we sketch event recollection.

1. Given an event of type T_{source}, one searches in the neocortex for a match accessible via an index in the hippocampus, a type T_{target} such that $T_{source} \sqsubset T_{target}$.
2. If one finds T_{target} such that $prob(match(T_{source}, T_{target}, \text{MaxQUD})) \geq \theta_{high}$, then $recall(T_{source})$ and $appraise(T_{target})$.
3. If all potential matches are evaluated as $\theta_{low} \geq prob(match(T_{source}, T_{target}, \text{MaxQUD}))$, then $\neg recall(T_{source})$.

Negative event recall has two consequences, an incrementation of *negative pleasantness* in Mood and the potential for clarification interaction (if to a co-present interlocutor or as a self-addressed question):

(15) a. $$\left[\begin{array}{l} \text{preconds} : \left[\begin{array}{l} \text{e} : \textit{RecType} \\ \text{c1} : \neg\text{Recall(spkr,e)} \end{array}\right] \\ \text{effects} : \left[\textit{NegativePleasantnessIncr}(\delta, \varepsilon)\right] \end{array}\right]$$

b. $$\left[\begin{array}{l} \text{preconds} : \left[\begin{array}{l} \text{e} : \textit{RecType} \\ \text{c1} : \neg\text{Recall(spkr,e)} \end{array}\right] \\ \text{effects} : \left[\begin{array}{l} \text{l-m.cont} : \textit{IllocProp} \\ \text{q} = \lambda P.P(\text{preconds.e}) : \textit{Question} \\ \text{c2}: \textit{Co-Propositional}(\text{l-m.content,q}) \end{array}\right] \end{array}\right]$$

On the model of the memory system sketched here, damage to the memory system can occur as follows:

- damage to the hippocampus: loss of event indices—some past experiences inaccessible, no way to create new event memories;
- damage to the perirhinal/anterolateral entorhinal cortex: fewer familiar individuals;
- damage to semantic memory: fewer means to talk about familiar individuals.

We summarize the basic structure of memory sketched here:

(16) Memory = $$\left[\begin{array}{l} \text{Episodic} : \left[\text{Conversational} : \textit{list}(\text{LDGBType})\right] \\ \text{HC-indices} : \textit{set}(\langle \text{n} : \mathbb{N}, \text{e} : \textit{RecType} \rangle) \\ \text{Entities} : \textit{set}(\textit{RecType}) \\ \text{Sem-mem} : \textit{set}(\textit{RecType}) \end{array}\right]$$

5 Discussion of Initial Examples

We can now return to reconsider the data from section 2.

Example (1) Initially we have a visual percept that includes several individuals; (in a tv size version of this scene) Putin is retrieved from entity memory, and retrieved from semantic memory is the fact that Putin is a leader and the topos 'leaders should wear formal clothes'.[13] The incongruity between the visual scene and the topos triggers the initial laugh. This leads to a pleasantness increment and the creation of a hippocampal index for the interaction and for the perceived visual scene. The interaction a year later involves successful recollection which can unify either on the index for the visual scene or for the conversational interaction. Whichever event is recalled is reappraised, so new potential for laughter.

Example (2) Originally Erdös had met Eliott Mendelson, who told him where he was from. This made EM and Vancouver familiar entities for Erdös, as well as updating his semantic memory in this respect. Due to Erdös's facial agnosia, when he encountered Eliott Mendelson, he was not (visually) familiar, which triggers the initial identity question. The answer to this question reveals the conceptually familiar entity Vancouver, which pattern completes to Elliot Mendelson, hence his deduction.

Example (3) In this case Carol's initial interaction is interrupted, which leads to the initial interaction being imperfectly recalled, perhaps via the mechanism proposed by Manohar et al. (2019) (viz. plasticity of synaptic weights; cf. section 3.4.1) and licensing the clarification interaction.

Examples (4) and (5) In both cases we have damaged semantic memories; the failed recollection

[13] Whether topoi live in semantic memory or in some more procedural section of memory we will not consider now.

licenses laughter in both cases, triggered by social incongruity the dementia sufferers are still aware of; the repeated recall failures take their toll in the depression exhibited in (5).

Example (6) This simply illustrates that successful recall triggers appraisal of the recalled event, with the consequent signals (laughter/crying) this can give rise to.

6 Conclusions and Future Work

In this paper we have sketched in rough outline a potential construal of certain aspects of dialogue context in terms of brain networks. We have suggested that this is the most parsimonious answer to the question of how to construe what dialogue contexts are in a way that directly captures memory fragility. This is, in turn, we have argued, pervasively present in interaction and needs to be integrated in accounts of dialogue coherence. At the same time we emphasize that the aim is not to replace computational theories of dialogue, which need to specify interaction in high level terms; the aim is to ensure bi-directional communication between such theories and theories formulated at the algorithmic and impementational levels of brain structures. While the roughness of our sketch is in no doubt, we believe that providing a dialogue–oriented semantics to models coming from neuropsychological research into memory has the potential of pushing such research to address spontaneous dialogue, which is an important aim.

References

Alan Baddeley. 1988. Cognitive psychology and human memory. *Trends in neurosciences*, 11(4):176–181.

Alan Baddeley. 2012. Working memory: Theories, models, and controversies. *Annual Review of Psychology*, 63:1–29.

Emmanuel J. Barbeau, Jérémie Pariente, Olivier Felician, and Michele Puel. 2011. Visual recognition memory: A double anatomo-functional dissociation. *Hippocampus*, 21(9):929–934.

Jon Barwise and John Perry. 1983. *Situations and Attitudes*. Bradford Books. MIT Press, Cambridge.

Christine Bastin, Gabriel Besson, Jessica Simon, Emma Delhaye, Marie Geurten, Sylvie Willems, and Eric Salmon. 2019. An integrative memory model of recollection and familiarity to understand memory deficits. *Behavioral and Brain Sciences*, 42.

James T. Becker, François Boiler, Oscar L. Lopez, Judith Saxton, and Karen L. McGonigle. 1994. The natural history of Alzheimer's disease: Description of study cohort and accuracy of diagnosis. *Archives of Neurology*, 51(6):585–594.

Sophia Bennett and Alan J. Thomas. 2014. Depression and dementia: cause, consequence or coincidence? *Maturitas*, 79(2):184–190. PMID: 24931304.

Ellen Breitholtz. 2020. *Enthymemes and Topoi in Dialogue*. Number 41 in Current Research in the Semantics/Pragmatics Interface. Brill, Leiden and Boston.

Steven L. Bressler and Vinod Menon. 2010. Large-scale brain networks in cognition: Emerging methods and principles. *Trends in Cognitive Sciences*, 14(6):277–290.

Donald T Campbell. 1974. 'downward causation'in hierarchically organised biological systems. In *Studies in the Philosophy of Biology*, pages 179–186. Springer.

Greg N. Carlson. 1977. *Reference to kinds in English*. Ph.D. thesis, University of Massachusetts.

Dennis Chan, Nick C. Fox, Rachael I. Scahill, William R. Crum, Jennifer L. Whitwell, Guy Leschziner, Alex M. Rossor, John M. Stevens, Lisa Cipolotti, and Martin N. Rossor. 2001. Patterns of temporal lobe atrophy in semantic dementia and Alzheimer's disease. *Annals of neurology*, 49(4):433–442.

Robin Cooper. 2022. *From perception to communication: An analysis of meaning and action using a theory of types with records (TTR)*. Oxford University Press.

Robin Cooper and Jonathan Ginzburg. 2015. Type theory with records for natural language semantics. In Shalom Lappin and Chris Fox, editors, *The Handbook of Contemporary Semantic Theory*, 2 edition, chapter 12, pages 375–407. Wiley-Blackwell, Oxford, UK.

Roxanne DePaul. 2017. DementiaBank English PPA DePaul Corpus. doi:10.21415/T5ZH5T.

Jonathan Erez, Rhodri Cusack, William Kendall, and Morgan D. Barense. 2016. Conjunctive Coding of Complex Object Features. *Cerebral Cortex*, 26(5):2271–2282.

Guido Gainotti. 2019. What face familiarity feelings say about the lateralization of specific entities within the core system. *Behavioral and Brain Sciences*, 42.

Jonathan Ginzburg. 2012. *The Interactive Stance: Meaning for Conversation*. Oxford University Press, Oxford, UK.

Jonathan Ginzburg and Andy Lücking. 2020. On laughter and forgetting and reconversing: A neurologically-inspired model of conversational context. In *Proceedings of the 24th Workshop on the Semantics and Pragmatics of Dialogue*, SemDial/WatchDial.

Jonathan Ginzburg, Chiara Mazzocconi, and Ye Tian. 2020. Laughter as language. *Glossa: a journal of general linguistics*, 5(1).

Mehrshad Golesorkhi, Javier Gomez-Pilar, Federico Zilio, Nareg Berberian, Annemarie Wolff, Mustapha C. E. Yagoub, and Georg Northoff. 2021. The brain and its time: intrinsic neural timescales are key for input processing. *Communications Biology*, 4.

Daniel L. Greenberg and Mieke Verfaellie. 2010. Interdependence of episodic and semantic memory: Evidence from neuropsychology. *Journal of the International Neuropsychological society*, 16(5):748–753.

Peter Hagoort. 2020. The meaning-making mechanism(s) behind the eyes and between the ears. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 375(1791).

Uri Hasson, Janice Chen, and Christopher J. Honey. 2015. Hierarchical process memory: Memory as an integral component of information processing. *Trends in Cognitive Sciences*, 19(6):304–313.

Hans Kamp. 1981. A theory of truth and semantic representation. In Jeroen Groenendijk, editor, *Formal Methods in Semantics*. Amsterdam Centre for Mathematics.

B. A. Kent, M. Hvoslef-Eide, L. M. Saksida, and T. J. Bussey. 2016. The representational–hierarchical view of pattern separation: Not just hippocampus, not just space, not just memory? *Neurobiology of Learning and Memory*, 129:99–106. Pattern Separation and Pattern Completion in the Hippocampal System.

Stefan J. Kiebel, Jean Daunizeau, and Karl J. Friston. 2008. A hierarchy of time-scales and the brain. *PLOS Computational Biology*, 4(11):1–12.

Walter Kintsch and Elizabeth Bates. 1977. Recognition memory for statements from a classroom lecture. *Journal of Experimental Psychology: Human Learning and Memory*, 3(2):150.

Staffan Larsson. 2002. *Issue based Dialogue Management*. Ph.D. thesis, Gothenburg University.

Andy Lücking. 2016. Modeling co-verbal gesture perception in type theory with records. In *Proceedings of the 2016 Federated Conference on Computer Science and Information Systems*, pages 383–392.

Brian MacWhinney. 2000. *The CHILDES Project: Tools for analyzing talk*, 3 edition. Lawrence Erlbaum Associates, Mahwah, NJ.

Sanjay G. Manohar, Nahid Zokaei, Sean J. Fallon, Tim P. Vogels, and Masud Husain. 2019. Neural mechanisms of attending to items in working memory. *Neuroscience and Biobehavioral Reviews*, 101:1–12.

Stephen Maren and Gregory J. Quirk. 2004. Neuronal signalling of fear memory. *Nature Reviews Neuroscience*, 5:844–852.

Chris B. Martin, Danielle Douglas, Rachel N. Newsome, Louisa L. Y. Man, and Morgan D. Barense. 2018. Integrative and distinctive coding of visual and conceptual object features in the ventral visual stream. *eLife*, 7:e31873.

Richard Montague. 1974. Pragmatics. In Richmond Thomason, editor, *Formal Philosophy*. Yale UP, New Haven.

Chi T. Ngo, Sebastian Michelmann, Ingrid R. Olson, and Nora S. Newcombe. 2021. Pattern separation and pattern completion: Behaviorally separable processes? *Mem Cognit*, 49(1):193–205.

Elizabeth A. Phelps. 2004. Human emotion and memory: interactions of the amygdala and hippocampal complex. *Current Opinion in Neurobiology*, 14(2):198–202.

David Poeppel and David Embick. 2005. Defining the relation between linguistics and neuroscience. In A. Cutler, editor, *Twenty-first century psycholinguistics: Four cornerstones*. Lawrence Erlbaum.

Blake A. Richards and Paul W. Frankland. 2017. The persistence and transience of memory. *Neuron*, 94(6):1071–1084.

Craige Roberts. 1996. Information structure in discourse: Towards an integrated formal theory of pragmatics. *Working Papers in Linguistics-Ohio State University Department of Linguistics*, pages 91–136. Reprinted in Semantics and Pragmatics, 2012.

Edmund T. Rolls. 2016. Pattern separation, completion, and categorisation in the hippocampus and neocortex. *Neurobiology of Learning and Memory*, 129:4–28.

James A. Russell. 2003. Core affect and the psychological construction of emotion. *Psychological review*, 110(1):145.

Talya Sadeh and Yoni Pertzov. 2020. Scale-invariant characteristics of forgetting: Toward a unifying account of hippocampal forgetting across short and long timescales. *Journal of Cognitive Neuroscience*, 32(3):386–402.

William Beecher Scoville and Brenda Milner. 1957. Loss of recent memory after bilateral hippocampal lesions. *Journal of Neurology, Neurosurgery & Psychiatry*, 20(1):11–21.

Tali Sharot and Andrew P. Yonelinas. 2008. Differential time-dependent effects of emotion on recollective experience and memory for contextual information. *Cognition*, 106(1):538–547.

Timothy J. Teyler and Jerry W. Rudy. 2007. The hippocampal indexing theory and episodic memory: updating the index. *Hippocampus*, 17(12):1158–1169.

Endel Tulving. 1972. Episodic and semantic memory. In E. Tulving and W. Donaldson, editors, *Organization of memory*. Academic Press, New York.

Bruce W. A. Whittlesea and Lisa D. Williams. 2000. The source of feelings of familiarity: the discrepancy-attribution hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26(3):547.

Andrew P. Yonelinas, Charan Ranganath, Arne D. Ekstrom, and Brian J. Wiltgen. 2019. A contextual binding theory of episodic memory: systems consolidation reconsidered. *Nature Reviews Neuroscience*, 20(6):364–375.

Andrew P. Yonelinas and Maureen Ritchey. 2015. The slow forgetting of emotional episodic memories: an emotional binding account. *Trends in Cognitive Sciences*, 19(5):259–267.