1 **Title:**
2 Cognitive neurorobotics and self in the shared world, a focused review of ongoing
3 research
4
5 Jun Tani
6 Cognitive Neurorobotics Research Unit,
7 Okinawa Institute of Science and Technology (OIST), Okinawa, Japan
8 tani1216jp@gmail.com
9
10 Jeffrey White (cor. au.)
11 Cognitive Neurorobotics Research Unit,
12 Okinawa Institute of Science and Technology (OIST), Okinawa, Japan
13 jeffreywhitephd@gmail.com
14
15

16 **Title:**

17

18 Cognitive neurorobotics and self in the shared world, a focused review of ongoing
19 research

20

21

22 **Abstract**

23

24 Through brain-inspired modeling studies, cognitive neurorobotics aims to resolve
25 dynamics essential to different emergent phenomena at the level of embodied agency in
26 an object environment shared with human beings. This paper is a review of ongoing
27 research focusing on model dynamics associated with human self-consciousness. It
28 introduces the free energy principle and active inference in terms of Bayesian theory
29 and predictive coding, then discusses how directed inquiry employing analogous models
30 may bring us closer to representing the sense of self in cognitive neurorobots. The first
31 section quickly locates cognitive neurorobotics in the broad field of computational
32 cognitive modeling. The second section introduces principles according to which
33 cognition may be formalized, and reviews cognitive neurorobotics experiments
34 employing such formalizations. The third section interprets the results of these and other
35 experiments in the context of different senses of self, both "minimal" and "narrative"
36 self. The fourth section considers model validity and discusses what we may expect
37 ongoing cognitive neurorobotics studies to contribute to scientific explanation of
38 cognitive phenomena including the senses of minimal and narrative self.

39

47

48

49

50

1. Prospectus

The following paper proceeds in four movements. The first very briefly introduces cognitive neurorobotics (CNR). The second section reviews a series of humanoid CNR experiments designed to elicit specific cognitive phenomena in emergent dynamics, and the third section interprets this review in the context of minimal and narrative self. The fourth section emphasizes the potential for CNR studies of this sort to contribute to inquiry into embodied cognition and mind.

1.1 Introducing Cognitive Neurorobotics

The main motivation of cognitive neurorobotics (CNR) is to elucidate essential mechanisms underlying embodied cognition through synthesis of analogous dynamics in various robotics experiments. CNR calls on diverse interdisciplinary knowledge including from fields of cognitive science, psychology, ethology, neuroscience, complex systems, AI & deep learning, artificial life and many others. Primarily, CNR can be considered as a marriage of two research fields. One is cognitive robotics (e.g., Levesque and Lakemeyer, 2008) which aims to develop human level intelligence in robots using a rather conventional symbolism approach, and the other is neurorobotics which puts more emphasis on the realization of adaptive behaviors of biological systems using neuroscience inspired models or neuromorphic schemes.

Initially, cognitive robotics studies proceeded with a strong conviction that formal logical descriptions of the world and rational computation for reasoning, planning, and inference could provide for human-level cognitive competency in artificial agents including robots. However, such early expectations were betrayed by the results of projects such as SRI's SHAKEY (Nilsson, 1984) which demonstrated the problems associated with the rigors of applying formal logic to real robots. Part of the reason for the trouble is that representation of the world using symbols as arbitrary tokens cannot be grounded smoothly with real world phenomena which are fundamentally given to experience in terms of continuous sensory-motor patterns. This is the point of Harnad's (1990) famous symbol grounding problem.

Already by the end of the 1980s, a paradigm shift had been taking place in AI and robotics research with the introduction of behavior-based robotics by Rodney Brooks (Brooks, 1990; 1991). Brooks considered that even simple insect-like robots can exhibit

87 extremely complex and intelligent behaviors by establishing a direct coupling between a
88 robot's reflex-type controllers and sensations from the environment. His seminal paper,
89 "Intelligence without representation" (Brooks, 1991) represents this thought -- no
90 representation and thus no grounding problem.
91
92 Neurorobotics in general is a growing field of especially fruitful inquiry employing
93 biological system inspired algorithms in a range of applications, from prosthetics with
94 brain-machine-interface technologies (Millan et al., 2010; Moxon and Foffani, 2015) to
95 independently embodied robots with autonomous locomotion, learning, memory, value
96 and action selection systems (Doya et al., 2002; Kuniyoshi and Sangawa, 2005; Kaplan,
97 2008; Krichmar, 2018). The degree of biological precision in selection of neuronal and
98 kinematic models depends on the degree of realism required to represent target
99 phenomena. Extreme realism is represented in the Human Brain Project (HBP).
100 Neurorobotics is considered a "strategic pillar" of the HBP through which biologically
101 inspired algorithms representing levels of organization from molecular mechanism to
102 modular function to unified cognitive architecture can be tested in simulation and then
103 deployed in physically embodied robots sharing physical space with human beings
104 (Knoll and Gewaltig, 2016). In general, however, more specific studies replicate focal
105 operations in fine detail while rendering other aspects more abstract. Recognizing the
106 impossibility of analyzing all levels of activity simultaneously for instance, Krichmar
107 and Edelman (2002) focus on how cortical and subcortical levels interact in real-time
108 using a relatively simple embodied robot, Darwin VII.
109
110 There has been a group of researchers who have emphasized advantages of studying so-
111 called "minimum cognition". These researchers have focused on phenomena emerging
112 during system-level interaction with the environment using relatively simple neuronal
113 adaptive controllers (e.g., Beer, 2000; Nolfi and Floreano, 2000; Iizuka and Di Paolo,
114 2007; Froese and Ziemke, 2009; Silberstein and Chemero, 2013; Barandiaran and
115 Chemero, 2006). Beer (2000) viewed an agent's nervous system, its body and its
116 environment as coupled dynamical systems. By focusing on the unfolding trajectory of
117 the agent's system state as shaped by both forces internal to the agent and external from
118 the environment, he attempts to extract the essential dynamic structures accounting for
119 minimal cognition. Since most neurorobotics studies inherit the aforementioned
120 thoughts of the behavior-based robotics, current research tends to stay close to the
121 realization of biologically plausible adaptive behavior, focusing on sensory-motor level

122 processing and hesitating to explore mechanisms associated with higher cognition in
123 human beings.
124
125 On the other hand, one of the main motivations of CNR studies is to consider possible
126 principles, algorithms, and implementation designs which can bridge the gap between
127 higher cognition and the lower sensory-motor processing of robots. Hybrid models (e.g.
128 Sun, 2002) attempt to combine these two levels, extracting symbolized rules and
129 associations at one level from sensory-motor patterns of activity at another (cf.
130 Kotseruba and Tsotsos, 2018). However, such an enterprise may suffer again from the
131 symbol grounding problem (Harnad, 1990) since these two levels do not share the same
132 metric space required for the dense interactions between top-down and bottom-up
133 processes that are associated with subjective experience including the sense of self.
134
135 Recently, deep learning schemes in robots show promise in attacking this problem. It
136 has been known that various types of deep learning networks can develop hierarchical
137 information processing in collective spatio-temporal activities of the neural units
138 through end-to-end learning of sensory-motor patterns. Actually, such trials have been
139 conducted by various research groups including the authors' (Yamashita and Tani,
140 2008; Levine et al., 2016; Yamada et al., 2016; Heinrich and Wermter, 2018).
141 Developmental robotics (Meta et al., 2008; Asada et al., 2009; Cangelosi and
142 Schlesinger, 2015) is another indispensable approach to address this problem. In
143 developmental robotics, cognitive competencies of artificial agents or robots develop
144 gradually, supported by human tutors, with scaffolding from one level to next level
145 according to fundamental theories in child development (e.g., Piaget, 1953; Vygotsky et
146 al., 1962).
147
148 Some CNR research attempts to expose connections between the phenomenology of
149 subjective experience and embodied sensory-motor processing. Holland (2007)
150 conjectured that building human-like bodies for robots and developing internal models
151 for predicting body dynamics is essential for developing "machine consciousness".
152 Prescott and Camilleri (2019) considers that the sense of self can be characterized as a
153 transient process, analogous to Tani's (1998) consideration that self becomes an object
154 of consciousness when prediction error for the actional outcome momentarily increases
155 (as will be described later, in section 3). Lanillos and colleagues (2017) proposed that
156 prediction error generated by a body's forward model should come with the sense of
157 what they call "enactive self" in the differentiation between inbody and other sources.

158 And, Hafner and colleagues (Lang et al., 2018; Schillaci et al., 2016) conducted a set of
159 robotics experiments also examining the sense of agency wherein they observed
160 attenuation of sensory inputs to self-generated movements (in terms of prediction error
161 minimization) but not to those of others (as these were unpredictable). Keeping this
162 research in mind, it is becoming more crucial that the problems of cognition and of
163 subjective experience should be investigated inseparably in order to gain a better
164 understanding of the minds of both humans and artifacts.
165
166 The CNR experiments reviewed in the next section aim to uncover such structural
167 dynamics by using an approach analogous to the free energy minimization principle
168 (FEP) proposed by Friston (2005). The FEP approach can be interpreted in terms of
169 Marr's three levels (Marr, 1982) wherein first the computational level might be
170 represented by the FEP itself, in which the goal of computation is minimizing the free
171 energy. Second, the representation and algorithmic level might be represented by the
172 schemes of predictive coding (Rao and Ballard, 1999; Friston, 2005; Clark, 2015) and
173 active inference (Friston et al., 2011; Hohwy, 2013; Clark, 2015). Finally, the
174 implementation level might be represented by neurophysiology in biological brains or
175 artificial neural network programs put in the robot's head. The next section starts with a
176 brief review of the FEP, predictive coding, and active inference, and then reviews a set
177 of CNR experiments conducted by Tani's group employing analogous principles. Their
178 validity as evidence for explanations in cognitive science is recalled in section 4.
179
180
181 **2. Models and CNR experiments**
182 First, we provide a brief introduction of the free energy minimization principle (FEP)
183 (Friston, 2005). Then, some neural network models developed by Tani's group similar
184 to this principle are introduced along with cognitive neurorobotics (CNR) experiments
185 using those models. These detailed descriptions of technical aspects may help to
186 understand how this and similar ongoing research may contribute to inquiry into
187 phenomena such as self, as explored in section 3 and as proposed in section 4.
188
189 **2.1 The free energy minimization principle**
190 The FEP states that any self-organizing system at dynamic equilibrium with its
191 environment must minimize its free energy in order to maintain this equilibrium and
192 thereby its organization in the face of otherwise disintegrative forces. The FEP is an
193 application of Bayesian theory. Bayes' general idea was that we can calculate the

194 probability of something being true (before we have evidence and rather than directly
195 testing for it), and to do so we may use what we have already observed to predict what
196 we should perceive next. By applying this principle to adaptation mechanisms in brains,
197 Friston considers that all essential cognitive mechanisms including perception, action
198 generation, and learning can be explained.
199
200 It may help to illustrate with an intuitive image of the FEP at work. Consider a
201 snowflake with wings, fluttering about. As long as it stays in a certain conditional zone
202 – not too turbulent, freezing air - then it continues fluttering about and is able to
203 maintain its unique structure. If it falls out of this zone, then its integrity is lost and it
204 dissipates. If its situation becomes too hot, then it melts, for example. In the simplest of
205 terms, the FEP (along with active inference, also introduced below) tells us that the
206 snowflake will do what it can to stay in this zone. As an application of Bayesian theory,
207 the FEP attempts to describe how biological brains update beliefs (or hypotheses, what
208 we think is the case) in light of new information, so that organisms can act on this new
209 information towards what they wish to be the case, like snowflakes countering gusts of
210 wind to stay in their comfort zones by anticipating from which direction the likeliest are
211 to come next.
212
213 Formally, the relationship between the likelihood of an observation after we have
214 experiential evidence, with what we thought that the likelihood might have been before
215 we had such evidence, is represented in a mathematical formula, Bayes' theorem.
216 Bayes' theorem updates prior probabilities (original hypotheses) with new evidence to
217 produce new (posterior, after the evidence is gathered) probabilities that are then useful
218 to guide the next iteration of action, e.g. the snowflake's next wing-flap. The FEP tells
219 us that biological brains do this to stay alive (like the snowflake does to maintain its
220 unique organization), so they (in general) perceive what they need to perceive to inform
221 beliefs that they need to believe to inform action that they need to enact in order to
222 maintain their integrity in the face of dissipative change. With its brain optimized
223 accordingly, an organism aims to minimize the difference between what it expects to
224 happen and what it perceives, especially concerning those parameters that bear on its
225 integrity; e.g. in order to stay in that comfort zone. In the end, it is this adaptive
226 updating of hypotheses in light of new information that is the focus of the CNR
227 experiments reviewed below, and it is the temporal hierarchy characterizing biological
228 brains grounding the anticipatory nature of the experience that their focal architectures
229 reflect.

230

231 In terms of these robotics experiments, Bayes' theorem is shown in formula 1. By
232 applying this formula, the robots are able to infer the hidden cause *z* of a sensory
233 observation *X*.

234

235
$$p(z|X) = \frac{p(X|z)p(z)}{p(X)} \tag{1}$$

236

237 P*(z/X)* is the posterior distribution of hidden cause *z* with given observation *X* and *p(X/z)*
238 is the likelihood which relates the sensory observation *X* to the hidden causes *Z*. *P(z)* is
239 the prior distribution of (probability density of) or belief in *z* as the latent cause before
240 observation *X*. *P(X)* is the marginal likelihood which is obtained by marginalizing
241 *p(X/z)* for all *z*. *P(X)* is obtained by considering the probabilities of *X* for all possible
242 hidden causes. There is a practical problem, however, in that the direct computation of
243 *p(X)* is often intractable with current methods. Simply too many possibilities must be
244 considered in practice.

245

246 Since *p(X)* is necessary for the computation of the posterior distribution *p(z/X)*, rather
247 than directly compute over all of these possibilities, the variational free energy approach
248 derives an approximation of *p(z/X)*. This scheme optimizes an auxiliary probability
249 distribution *q(z)*, referred to as the recognition density, in approximation of the true
250 posterior *p(z/X)* by minimizing the Kullback-Leibler divergence (KL divergence)
251 between the two (formula 2).

252
$$\boldsymbol{D_{KL}[q(z) \parallel p(z|X)]} = \int q(z)ln\frac{\boldsymbol{q(z)}}{\boldsymbol{p(z|X)}}\boldsymbol{dz}$$
253
$$= \boldsymbol{F + lnp(X)} \tag{2}$$

254

255 Free energy *F* is defined as

256

257
$$F = \int q(z)ln\frac{q(z)}{p(z,X)}dz \tag{3}$$

258

259 Since the marginal likelihood *p(X)* in the second term on the right-hand side in (2) is
260 independent of the recognition density *q(z)*, minimization of the KL divergence between
261 the recognition density and the true posterior can be achieved by minimization of the

262   free energy in (3) respect to *q(z)*. This makes *q(z)* as an adequate practical

263   approximation of the true posterior *p(z/X)*.

264

265   The free energy *F* to be minimized can thus be rewritten in a computationally tractable

266   form in terms of *q(z)*. In this equation, **φ** and **θ** are model parameters (formula 4):

267

268 $$F = -E_{q_\varphi(z)}[\ln p_\theta(X|z)] + D_{KL}[q_\varphi(z)||p(z)] \qquad (4)$$

269

270   The first term on the right side of the equation, the accuracy term, represents the

271   expectation of log-likelihood with respect to the approximate posterior, which

272   represents reconstruction of the sensory observation with the approximate posterior by

273   generative models. The second term, the complexity term, is represented by KL

274   divergence between the approximate posterior and the prior, which serves to regularize

275   the model according to prior expectation.

276

277   Free energy *F* can be minimized with respect to $q_\varphi(z)$ as

278

279 $$q_\varphi(z) = \arg\max F \qquad\qquad (5)$$

280

281   Here, we may put the preceding in terms of predictive coding (Rao and Ballard, 1999;

282   Friston, 2005; Clark, 2015). For a given sensory observation, the posterior inference of

283   the latent variable ("latent" means hidden, so this is what is hypothesized to be the

284   hidden cause) is carried out by means of minimizing the error between the sensation

285   expected by the generative model for the latent variable, and its observation under the

286   constraint of the prior distribution of that latent variable. The idea here again is to

287   minimize the difference between what was predicted, and what is perceived.

288

289   Free energy *F* integrated for the predicted future time period is the expected free energy

290   F^E of a given model. We aim to minimize this value in respect to an action *a* by

291   assuming a forward model that represents the likelihood of an action causing an

292   observation given the cause *z*, as shown in formulae (6) and (7).

293

294 $$F^E = -E_{q_\varphi(z)}[\ln p_\theta(X(a)|z)] \qquad\qquad (6)$$

295

296 $$a = \arg\max F^E \qquad\qquad (7)$$

297

298   Note that the expected free energy does not involve the KL divergence between the

299   posterior and the prior (as in formula 4) because KL divergence is independent of

300   sensation (evidence, X) and thus of action.

301

302   The preceding process represents action generation by active inference (Friston et al.,

303   2011; Hohwy, 2013; Clark, 2015) whereby actions are selected such that the expected

304   free energy is minimized.[1] This stands for action generation by active inference (Friston

305   et al., 2011; Hohwy, 2013; Clark, 2015) where actions are selected such that the

306   expected free energy is minimized. (Later we introduce a more recent update of active

307   inference (Kaplan and Friston, 2018) considering the epistemic value.) More intuitively,

308   to minimize the error between the expected or preferred sensory outcome and the actual

309   one, the actual one is modified to become closer to the preferred one by acting

310   adequately on the environment. Finally, perception by (5) and action generation by (7)

311   are carried out simultaneously for closing the loop between action and perception

312   (Baltieri and Buckley, 2017). Here, it is considered that the enactive cognition is a

313   continuing trial for minimizing the conflictive error between the top-down intention

314   projected from the latent variables and the bottom-up perception from the reality

315   through iterative acts of changing the external environment as well as modifying the

316   intention within (Tani, 2016).

317

318   **2.2 Recurrent neural network with parametric biases**

319   Tani and colleagues (Tani, 2003; Tani et al., 2004) developed the recurrent neural

320   network with parametric biases (RNNPB) independently from the FEP. Since that time,

321   the RNNPB has turned out to be one possible neural network implementation of the

322   FEP (with some simplification as detailed later). Although there have been variations of

323   the RNNPB specific to different applications, the following describes a typical version.

324

325   **(A) Model description:**

326   In the RNNPB, the following objective function for a time series of L time steps is

327   minimized (formula 8):

328

---

[1] It is interesting in this context to consider Bruineberg et al. (2018) who modeled foraging agents learning most efficient paths to prior-specified goals, "desire paths". These models extended the FEP and active inference to show that expectations are shaped by changes to the environment, resulting in reduced free energy for the agent-environment system.
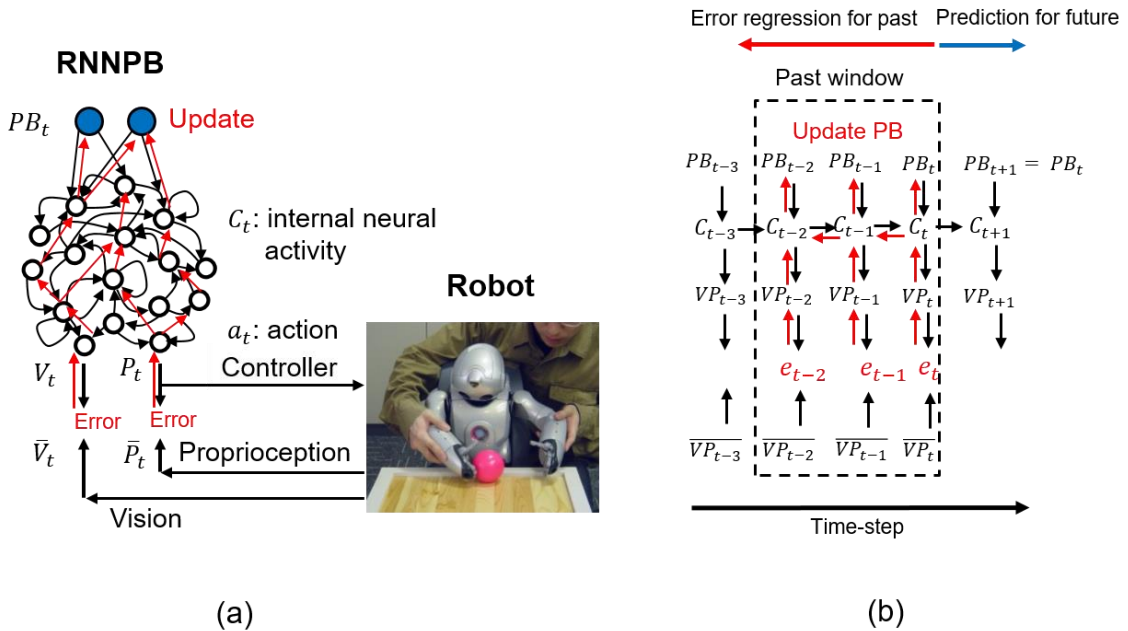
$$329 \quad Obj(pb_{1..T}, W) = \frac{1}{2} \sum_{t=1}^{L} \left(\frac{1}{N^o \sigma^2} \|\overline{X_t} - X_t(W)\|^2 + \frac{1}{N^{pb}} pb_t^2 \right) \quad (8)$$

330

331 In (8), $\overline{X_t}$ and $X_t(W)$ are the sensory observation and the prediction outputs for the

332 sensory observation as a function of the learning parameters W with their

333 dimensionality N^o. Sensation may consist of exteroception such as vision v_t and

334 proprioception p_t in a robotics application as shown in Figure 1 (a). In (8), $pb_{1..T}$ is a

335 latent variable of N^{pb} dimensions.

336

337 This objective function (8) becomes equal to the free energy defined in (4) with

(a)   (b)

338

339 Figure 1: A robot controlled by an RNNPB model. During movement, PB values are

340 updated by backpropagation in the direction of minimizing error (a); predicting the

341 future by inferring the most likely past with black and red arrows representing the

342 generative and error back-propagation processes during prediction and postdiction,

343 respectively (b).

344

345 assumptions[2] that the prior of the latent variable is represented by a unit Gaussian

346 distribution, the posterior by a Gaussian distribution with its mean pb_t and its standard

---

[2] In the interests of simplicity, we neglect here discussion of merits of variational Bayes models such as those representing the strength of belief or an estimate of precision in prediction. Recent developments

347 deviation of 1.0, and the sensory prediction outputs with a Gaussian distribution with its
348 mean X_t and its standard deviation σ. Then, this objective function can be minimized
349 with respect to $pb_{1..T}$ given fixed learning parameters W by the deterministic dynamics
350 described in the following difference equations (formulas 9a-9d):
351

$$
\begin{cases}
C_{t+1}^i = g\left(\sum_j w_{ij}^c C_t^j + \sum_h w_{ih}^{pb} PB_{t,n}^h + b_i^c\right) & (9a) \\[2ex]
pb_{t,n+1}^h = \alpha\left(-pb_{t,n}^h - \nabla E\left(pb_{t,n}^h\right)\right) + pb_{t,n}^h & (9b) \\[2ex]
PB_{t,n}^h = g\left(pb_{t,n}^h\right) & (9c) \\[2ex]
X_{t+1}^o = g\left(\sum_j w_{oj}^{out} C_t^j + b_o^{out}\right) & (9d)
\end{cases}
$$

353
354 In (9a), C_t represents the activation of internal units with recurrence at time step t.
355 $PB_{t,n}$ represents the outputs of sigmoid function $g()$ applied to the latent variable
356 $pb_{t,n}$ during the *nth* epoch iteration at time step t. The latent variable pb is inferred at
357 each time step through N epochs of iterations of the internal computation loop. Its value
358 is updated in the direction of minimizing prediction error in the output by following
359 $\nabla E\left(pb_{t,n}^h\right)$ which is the gradient of the mean square error E of the prediction outputs for
360 all time steps while also considering a unit decay effect of this value as represented in
361 the first term on the right-hand side of (9b). w^c, w^{pb}, w^{out}, b^c, and b^{out} are elements
362 of the learning parameters W.
363
364 The actual computation of (9) is carried out for N epochs iterated in the forward
365 computation of L time steps, with backward error regression computed for the same L
366 time steps using a past window storing all temporal variables from time step t-L+1 to
367 the current time step t as illustrated in Figure 1(b). In the forward computation, values
368 of the internal units as well as those of the prediction outputs from time step t-L+1 to
369 the next (anticipated future) time step t+1 are computed using the PB values updated for
370 each time step. In the backward error regression, the prediction error at each time step is
371 back-propagated through time (BPTT) (Werbos, 1990; Rumelhart et al, 1986) from the
372 current time step t to time step t-L+1 (at the onset of the past window). The latent

using RNN models by the author's group (Murata et al., 2015; Ahmadi and Tani, 2019) address these
issues.

373   variable $pb_{t-L+1..t}$ in the window is updated by using the error gradient information

374   obtained by BPTT at each time step according to (9b).

375

376   The predicted proprioception at time step $t+1$ is sent to the robot PID (proportional-

377   integral-derivative) controller as target joint angles for the next time step of the robot,

378   and the robot moves accordingly. For instance, actuators receive the motor commands *a*

379   generated by the PID controller, move according to these commands, and this

380   movement generates new visual and proprioceptive (bottom-up) sensations and their

381   (top-down) prediction errors at time step $t+1$ again with the aim to minimize these

382   errors. The PID controller thereby instantiates the idea of active inference since action *a*

383   is generated in order to minimize the error between the actual and the predicted joint

384   angles (proprioception) (Baltieri and Buckley, 2018).

385

386   Finally, the learning of the RNNPB can be carried out by minimizing objective function

387   (8) with respect to both the learning parameters W and the latent variable $pb_{1..T}$

388
$$W_{n+1} = -\beta \nabla E(W_n) \qquad (10)$$

389

390   *W* is updated at each epoch using the gradient information computed by BPTT with the

391   learning rate β.

392

393   **(B) Robotics experiment using RNNPB**

394   To evaluate the RNNPB in a robotics experiment, a humanoid robot was used which

395   consists of three subsystems: an onboard sensory processing module with a head-

396   mounted video camera, an RNNPB module running on an external computer, and an

397   onboard motor control module. So configured, the RNNPB could predict two types of

398   sensory inputs, proprioception in terms of joint angles in both arms of the robot and

399   visual features representing target object position (X-Y-Z) at each time step.

400

401   During ball manipulation experiments (Ito et al., 2006), human tutors trained this robot

402   to manipulate a ball in two different sequences: repetitively pushing the ball from left to

403   right and right to left, and repetitively grasping the ball at the center position, lifting it

404   up and then dropping it. After the training of the network wherein PB values adapted

405   differently for each action sequence, the robot was tested to generate these sequences

406   autonomously (without external help). During testing, the robot switched from one to

407   the other intermittently, with an example presented in Figure 2 (also see a video:

408   https://youtu.be/a_auIoksGN0.) In this instance, the robot initially pushed the ball from

409 left and right repeatedly until the ball bounced off of one hand too much, rolling to the

410 center position. This unexpected movement caused prediction error. To minimize this

411 error, the robot adopted the PB value for grasping and dropping the ball, and its

412 behavior changing accordingly. Here, we see that the top-down intention to act with the

413 ball, represented by the PB, shifted autonomously during iterative interaction with the

414 bottom-up percept of the ball position from one behavioral reparatory to another in the

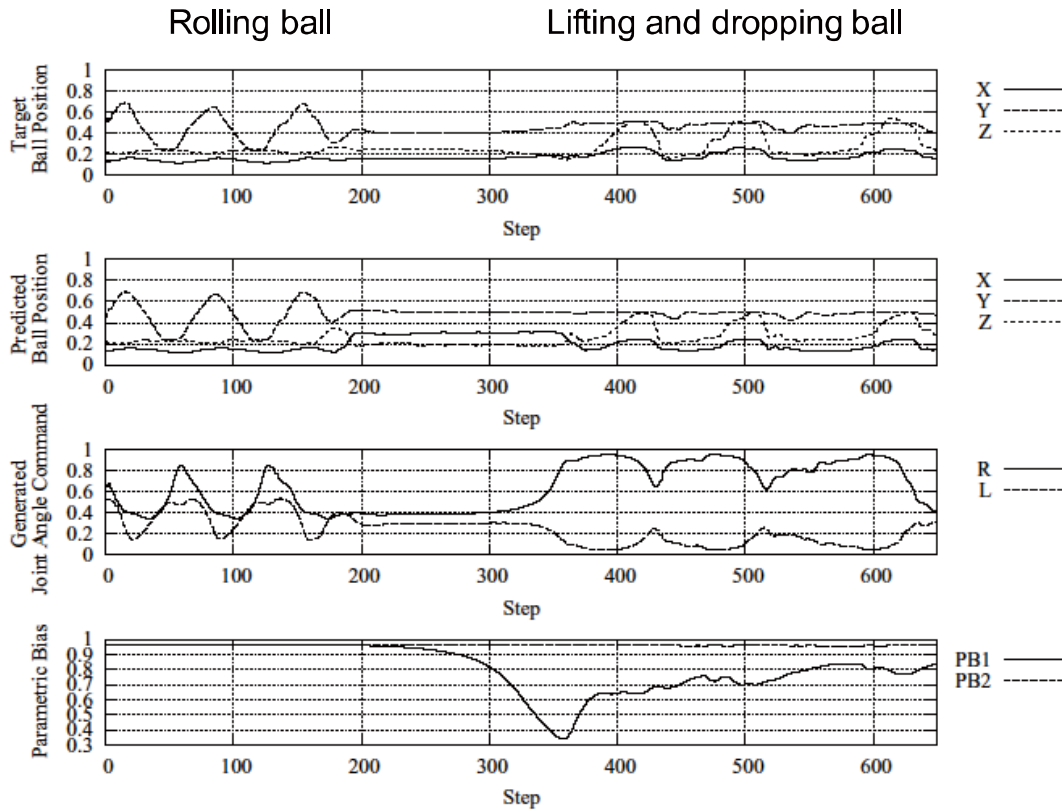415 course of minimizing the error.

416

417

418 Figure 2: Dynamic generation and switching of two learned ball handling behaviors.

419 Top row: measured ball position. Second and third rows: predicted ball position and

420 robot joint angles generated by the RNNPB, respectively. Bottom row: the PB as it

421 switches from one movement sequence to another. Redrawn from (Ito et al., 2006).

422

423 Importantly, the continuous sensory flow was segmented during the error minimization

424 process as the robot optimized coordination with the target object through autonomous

425 shifts in PB values. Shifting PB values served as bifurcation parameters to induce

426 transitions, effectively steering the system from one learned behavioral pattern to the

427 other. This same mechanism for the segmentation and chunking of the continuous

14

428   perceptual flow had been observed during an earlier RNNPB experiment on human-
429   robot interaction (Ito and Tani, 2004) wherein a robot and a human participant
430   attempted to synchronously imitate one another's primitive movement patterns using
431   prediction based on prior learning. With robot and human participant movements
432   synchronized according to one of the prior learned patterns, if the human participant
433   suddenly shifts the current movement pattern to another learned one, the
434   synchronization breaks down thereby generating prediction error bottom-up in the
435   RNNPB. In effort to minimize this error, the PB value is updated, and enactive
436   synchrony with the present pattern achieved. Looking at the data from these
437   experiments, segmentation of sensory flow from one habituated pattern (with
438   corresponding expectations) to another can also be observed.
439
440   Observation of these phenomena suggests a general mechanism for the segmentation
441   and chunking of the continuous sensory flow. Confirming a core tenet of predictive
442   coding, that cognition aims to minimize prediction error in the process of interacting
443   with changing environments, perceived error should be essential to mechanisms
444   underwriting the segmentation of the continuous bottom-up sensory flow during online
445   cognition in biological models, as well. Moreover, such mechanisms should provide for
446   the development of compositionality in cognition, subject of the next section, because
447   this competency requires segmentation of sensory flow into a set of reusable objects
448   which can be mentally manipulated for combinatory operations.
449
450   Here, it is intriguing to note that the error regression scheme for the past window in the
451   RNNPB could provide a possible mechanism for "postdiction" (Shimojo, 2014).
452   Postdiction is a process that is recognizable during perceptual phenomena wherein the
453   percept of a stimulus presented earlier is affected by another stimulus presented later.
454   Postdiction is apparent during various backward perception phenomena including
455   classic examples of backward masking (Raab, 1963) or the cutaneous rabbit illusion
456   (Geldard and Sherrick, 1972). Such phenomena may be explained by the error
457   regression mechanism assumed in the RNNPB as a model of predictive coding.
458
459   The cutaneous rabbit involves tactile stimuli (taps, small electric shocks) presented to a
460   human subject, typically on the forearm (due to this area's relatively poor spatial acuity
461   as mapped to the somatosensory cortex). In the simple case, three stimuli are presented
462   with equal temporal intervals between each. The cutaneous rabbit illusion appears when
463   these stimuli occur with very short durations between them (less than 300 milliseconds)

15

464 and when the first and the second are given in the same position, but with the third

465 given in a distant position (which may extend past the physical body, Miyazaki et al.,

466 2010). The subject mis-locates the second tap in the direction of the third tap, whereas

467 the subject will not mis-locate it if the third tap is given in the same position with the

468 first or second ones.

469

470 In backward masking, the consciousness of a target presented immediately before a

471 masking stimulus (typically something driving urgent attention, such as something

472 scary) can be suppressed, such that subjects are unable to report having seen the first

473 stimulus. This phenomenon may be intuitively explained given two assumptions, one

474 that the world as we experience it usually doesn't change so rapidly, and two that

475 optimal implicit internal models of such a world should operate according to the

476 expectation that a visual stimulus given a moment ago will be retained for a while. This

477 gives rise to the idea that there may be two pathways operative in backward masking,

478 one for normal operations over longer timespans, and one for surprising situations that

479 respond rapidly to changes in the environment. In case that a stimulus presenting minor

480 changes in the world is suddenly followed by a surprising stimulus presenting more

481 important changes in the world, an expectation error should be generated that

482 effectively diminishes the previous stimulus as the agent recenters activity around the

483 implications of the later stimulus. Thus, in the course of minimizing the error between

484 what is expected and what is sensed, the experience of the stimulus presented in the past

485 is "masked" by one that comes later.

486

487 The case of the cutaneous rabbit can be explained similarly. After the first and the

488 second (series of) taps are provided in the same position, an internal model implicitly

489 expects that the third should come in the same position after the same temporal interval.

490 However, when the third is presented in a distant position, an expectation error is

491 generated. To minimize the error, the experience of the second tap is relocated to the

492 midpoint between the first and the third tap positions (by means of regressing a linear

493 model predicting position and timing for succeeding stimuli). In sum, postdictive

494 phenomena can be explained in terms of inference and consequent rewriting of past

495 experience by means of the error minimization principle.[3]

496

---

[3] Here, we may understand "inference" specifically, as search for an optimal latent variable to minimize the error between the expected sensation and the actual one

2.3 Multiple timescale RNN for development of compositionality

The RNNPB experiments reviewed above demonstrate that cognitive phenomena such as segmentation of the undifferentiated perceptual flow into reusable chunks occurs by minimizing conflicts between bottom-up perception and top-down intention. This section briefly reviews experiments intended to expose the role of such dynamics in the composition of novel patterns in coordination with novel task environments before turning to how these fundamental dynamics may contribute to accounts of self.

Compositionality is the ability to combine parts into wholes, evident for example in the abilities to determine the meanings of sentences from the structured relations between constituent parts (Evans, 1982; Costello and Keane, 2000) and to enact diverse goal-directed actions by sequentially combining reusable primitives (Arbib, 1981; Pastra and Aloimonos, 2012). Yamashita and Tani (2008) proposed a predictive RNN model characterized by multiple time constraints at different levels, the multiple timescale recurrent neural network (MTRNN) to investigate how neural networks in biological brains develop compositionality and thereby generate novel actions. The MTRNN has been used to investigate various aspects related to development of compositionality including co-development of skills between human tutors and robots (Nishimoto and Tani, 2009), goal-directed creative compositions of primitives (Arie et al., 2009), cases analogous to schizophrenic pathology including the delusion of control (Yamashita and Tani, 2012), and imitative human-robot interactions (Hwang et al., 2019). Next, we review the MTRNN in greater detail.

(A) Model description:

The MTRNN has layers of RNNs each characterized by a different timescale constraint (see Figure 3(a) for a typical architecture). Neural activity in higher layers is slower with larger timescale constraints, while lower layers are faster with smaller timescale constraints. In the MTRNN, the following objective function for a time series of L time steps is minimized as analogous to (9) in the case of the RNNPB:

$$Obj(c_1, W) = \frac{1}{2} \sum_{t=1}^{L} \left(\frac{1}{N^o \sigma^2} \|\overline{X_t} - X_t(W)\|^2 + \frac{1}{N^c} c_1^2 \right) \qquad (11)$$

In (11), N^o and N^c represent the dimensions of the output and the internal units, respectively. (11) differs from (9) in that c_1 - which represents the potential value of

530   the internal units with recurrence at the initial time step in (9) - plays the role of latent

531   variable in (11).

532

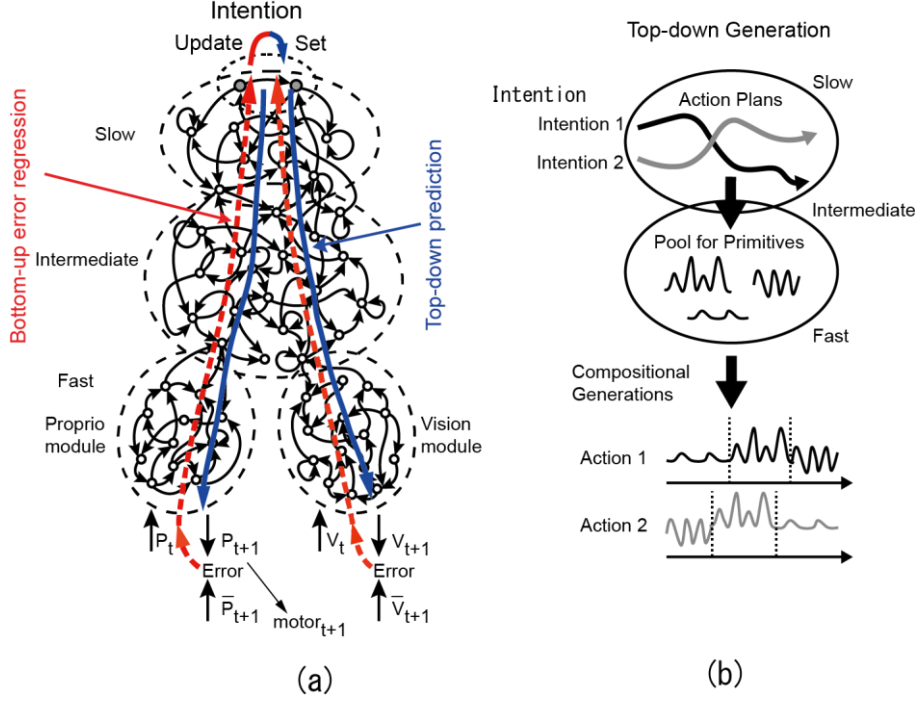533   (a)   (b)

534

535   Figure 3: (a) an example of an MTRNN architecture with 3-layers, and (b) illustration

536   of compositionality of different action plans in an MTRNN. Redrawn from (Tani,

537   2016).

538

539   Assuming fixed learning parameters of *W,* this objective function can be minimized

540   with respect to c_1 by way of the following difference equations (12a-12d):

541

542

$$\left\{ \begin{array}{ll} c_{t+1}^i = -\left(1 - \dfrac{1}{\tau^i}\right)c_t^i + \dfrac{1}{\tau^i}\sum_j w_{ij}^c C_t^j + b_i^c & for \ \ t \geq 1 \quad (12a) \\[3mm] c_{1,n+1}^i = \alpha\left(-c_{1,n}^i - \nabla E\left(c_{1,n}^i\right)\right) + \ c_{1,n}^i & (12b) \\[3mm] C_{t+1}^i = g\left(c_{t+1}^j\right) \qquad for \ \ t \geq 0 & (12c) \\[3mm] X_{t+1}^o = g\left(\sum_j w_{oj}^{out} C_{t+1}^j + b_o^{out}\right) & (12d) \end{array} \right.$$

543

544 In this series of equations, C_t represents activation of the internal units with recurrence
545 at time step t. τ^i is the time constant for each internal neural unit, and is different at
546 each layer (in comparison, the time constant for all internal units in the RNNPB was
547 assumed to be 1.0.) By assigning the higher layer with a larger value for the time
548 constant, the higher layer is dominated by slower dynamics; by assigning a smaller time
549 constant to the lower layer, it is dominated by faster dynamics. It is also noted that a
550 bottleneck type connectivity constraint was applied to the network wherein the internal
551 units in the higher layer were connected only with those in the intermediate layer which
552 were again connected with only those in the lower layer.
553
554 Here, we may note that the latent variables represented as the initial states of the internal
555 units c_1 determines the time development of the whole network, including later
556 prediction outputs, due to the sensitivity to initial conditions that is characteristic of the
557 deterministic RNN model. Therefore, for minimizing the reconstruction error for all
558 time steps, the optimal values of the initial states are inferred by using the gradient
559 information computed by BPTT through iterations of the internal computation loop
560 using the past window in a similar way with the RNNPB case.
561
562 **(B) Robotics experiment:**
563 A robotics experiment using an MTRNN in a task of developmental tutoring of
564 compositional object manipulation is briefly reviewed. During this experiment
565 (Nishimoto and Tani, 2009), a Sony QRIO (as in Figure 1) controlled by an MTRNN
566 was tutored on multiple task sequences each composed of different series of primitive
567 actions. For example, one tutored sequence proceeded as follows. The tutor moved both
568 of the robot's hands toward an object located at an arbitrary position on a table, then
569 using the hands grasped it, lifted it up and down a few times, and placed it back on the
570 table. Another sequence involved touching the object with the left and right hands in
571 turn, grasping the object, rotating it, and placing it back on the table. Sensors on the
572 robot delivered simplified visual features and proprioceptive information (as in
573 experiments described above). Tutoring proceeded gradually, i.e., the robot was tutored
574 on some tasks, then tested, and if performance was unsatisfactory, tutoring was repeated
575 (the corresponding video can be seen at https://youtu.be/n9NYcG8xlYs). It is important
576 to emphasize that tutors directly guided the robot's hands by feeling and correcting the
577 "intent" of the robot with their own hands. After training, the robot was able to reliably
578 perform all task sequences successfully and layer-specific neural regions were analyzed.

19

579 It was found that each layer played a different role in action compositionality (see
580 Figure 3 (b)).
581
582 Yamashita and Tani (2008) speculated that the multiple timescale property imposed on
583 network dynamics resulted in the emergence of a functional hierarchy in which the
584 higher layer generated different slowly changing neural activation patterns
585 corresponding to a scenario or plan for each task sequence, whereas the lower layer
586 developed precise skillful control for each behavior primitive. Recalling experiments
587 involving the RNNPB, the slowly changing neural activity from the higher layer served
588 as a source of bifurcation parameters, steering activity in the lower layer, while the
589 faster lower layer learned to develop a set of behavior primitives, e.g. grasping, lifting
590 up, or moving the object left and right, in complementary ways. After the development
591 of these dynamics through learning, the MTRNN-driven robot became able to generate
592 different ways of combining behavior primitives. Yamashita and Tani (2008)
593 interpreted these results by hypothesizing that the robot developed compositionality in
594 generating actions as its functional hierarchy self-organized. In order to test this
595 interpretation, they manipulated the timescale parameters of the MTRNN in order to
596 explore the role of the multiple timescales in structuring cognitive dynamics.
597 Compositional representations (layer-specific stable structural dynamics) could not be
598 developed when all layers shared the same time constant. The same results were
599 repeatedly confirmed in other robotics experiments conducted under more complex
600 conditions, for example while using pixel level vision (Hwang et al., 2018).
601

602 **2.4 Goal-directed planning**
603 With this understanding of compositionality as achieved through the segmentation of
604 fluid experience, predictive coding and active inference (see section 2.1) can be applied
605 also to the problem of the goal-directed planning of action sequences by robots.
606
607 Tani (1996) conducted experiments on the goal-directed navigation of a mobile robot,
608 Yamabico, which was equipped with a range sensor using a single layer RNN.
609 Yamabico was developed with a lower-level automatic control scheme using range
610 sensors which can perceive range images of 24 angular directions covering the front of
611 the body to travel smoothly in a collision-free manner in a workspace. Basically, it
612 moves toward the largest open space in a forward direction by maneuvering between
613 obstacles on its left and right sides. When a new open space appears, a decision is made

614   on whether to pursue the current open space direction or to branch to the new one. This
615   branching decision is made by the RNN in the higher cognitive level as described next.
616
617   In the experiment, Yamabico explored the obstacle workspace under collision-free
618   maneuvering control during daytime for the purpose of gathering sensory-action data.
619   When it encountered a branching point, a branching decision was made arbitrarily (with
620   an equal chance for either option) by the experimenter. At this time, the sensory inputs
621   in terms of the range image, the travel distance (as indicated on the odometer) from the
622   previous branching point to the current one, and the action in terms of the branching
623   decision, were recorded. Yamabico explored the environment experiencing around 200
624   successive branches until its battery was depleted. This resulted in a sampling of a
625   sensory-action sequence of around 200 branching steps. During nighttime while the
626   battery was charging, the RNN was trained in the form of the forward model (Miall and
627   Wolpert, 1996; Kawato, 1999) so that it could predict the sensory input at a next
628   branching step $X_{\tau+1}$ when presented with current sensory input X_τ and branching
629   action $a_{\tau+1}$ by developing an adequate dynamic structure for the latent state transition
630   from Z_τ to $Z_{\tau+1}$ based on the branching scheme using the sampled sensory-action
631   sequence.
632
633   After this training, a test of goal-directed planning was conducted through the following
634   procedure. First, Yamabico traveled for several steps by randomly branching out from
635   an arbitrary position for the purpose of inferring the latent state by way of the observed
636   sensory sequence.[4] Several branching steps of travel were necessary because this
637   navigation problem involves the sensory aliasing problem; the current sensory inputs
638   cannot uniquely identify the current latent state. Then, after inference of the latent state
639   in the current branching step, Yamabico generated goal-directed planning under the
640   constraint of minimum travel distance to reach a branching point specified as a goal by
641   its expected sensory inputs.
642
643   Analogous with active inference as in (7), an optimal action sequence minimizing the
644   error between the preferred and the predicted sensory outcomes was identified,

---

[4] This process can be related to the process of perception by predictive coding shown in (5). However, in this case, the latent variable is not inferred by minimizing prediction error. Instead, the latent variable is updated by means of the entrained (learned) sensory input sequence which turns out to minimize the prediction error.

645   specifically minimizing the error between the goal range image (image of the goal as it
646   is ideally achieved) and the predicted one in the distal (final) goal step (given the
647   current action plan) while also minimizing the predicted travel distance at each
648   branching step. This can be carried out by means of BPTT applied to the trained RNN
649   as illustrated in Figure 4. In the test experiment, a set of possible action sequence plans
650   was searched (including sub-optimal ones) through iterative computation. Generated
651   action plans were actually executed by the robot.
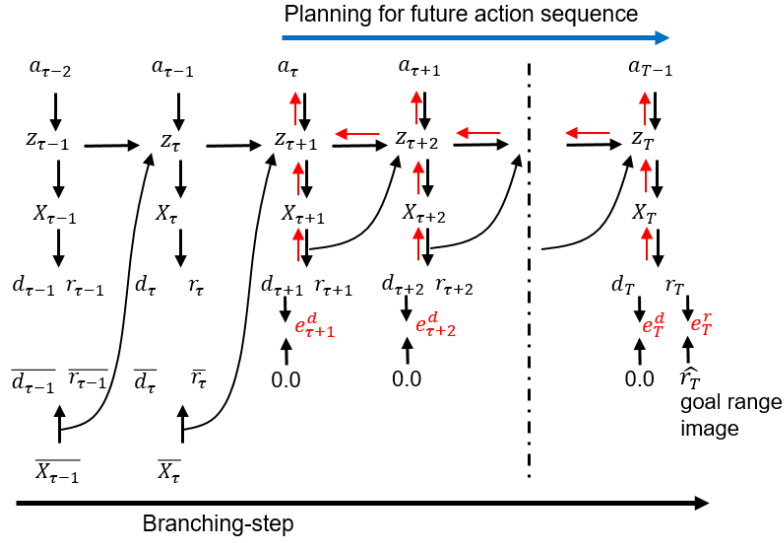652

653

654   Figure 4: Goal-directed planning using RNN. The future action sequence $a_{\tau..T-1}$ is
655   optimized for minimizing the expected travel distance $d_{\tau+1..T}$ at each branching step as
656   well as the error between the goal range image at the distal step $\hat{r}_\tau$ and its expected
657   value r_τ. The both errors for the travel distance at each step $e^d_{\tau+1..T}$ and the goal range
658   image at the distal step e^r_τ are back-propagated through steps.

659

660   Results of goal-directed plan generation are shown in Figure 5. Figure 5 (a) shows the
661   designated starting position and the goal position. Figure 5 (b)-(d) shows the actual
662   trajectories generated by executing 3 different action sequence plans, with Figure 5 (b)
663   showing the optimal trajectory that minimizes the travel distance, and Figure 5 (c-d)
664   showing suboptimal plans. Especially, it is noted that the robot had never enacted the
665   trajectory in Figure 5 (d) during the exploration phase before learning. This result
666   implies that Yamabico became capable of mentally imaging novel compositions of
667   experienced parts of trajectories by extracting the hidden structure of the environment
668   through consolidative learning of diverse sensory-action trajectories sampled.

669

670   To examine the internal structure developed by the RNN, a phase space analysis was

671   conducted. The RNN generates mental simulations of thousands of consecutive steps of

672   random branching sequences. Figure 5 (e) shows the phase space plot wherein each

673   point represents an internal state (projected in 2-dimensional space) when visiting a

674   particular branching point. It can be seen that the points are all clustered into a set of

675   segments. It was found that each segment corresponds to a particular branching point,

676   and that the points inside each segment form a Cantor set-like assembly. These

677   observations inform us how the compositionality is represented in the internal state of

678   the RNN. When the robot moves from one branching point to another by choosing a

679   particular branching action, the internal state transits from one segment to another.

680   Therefore, a graph-like state transition structure can be seen by considering those

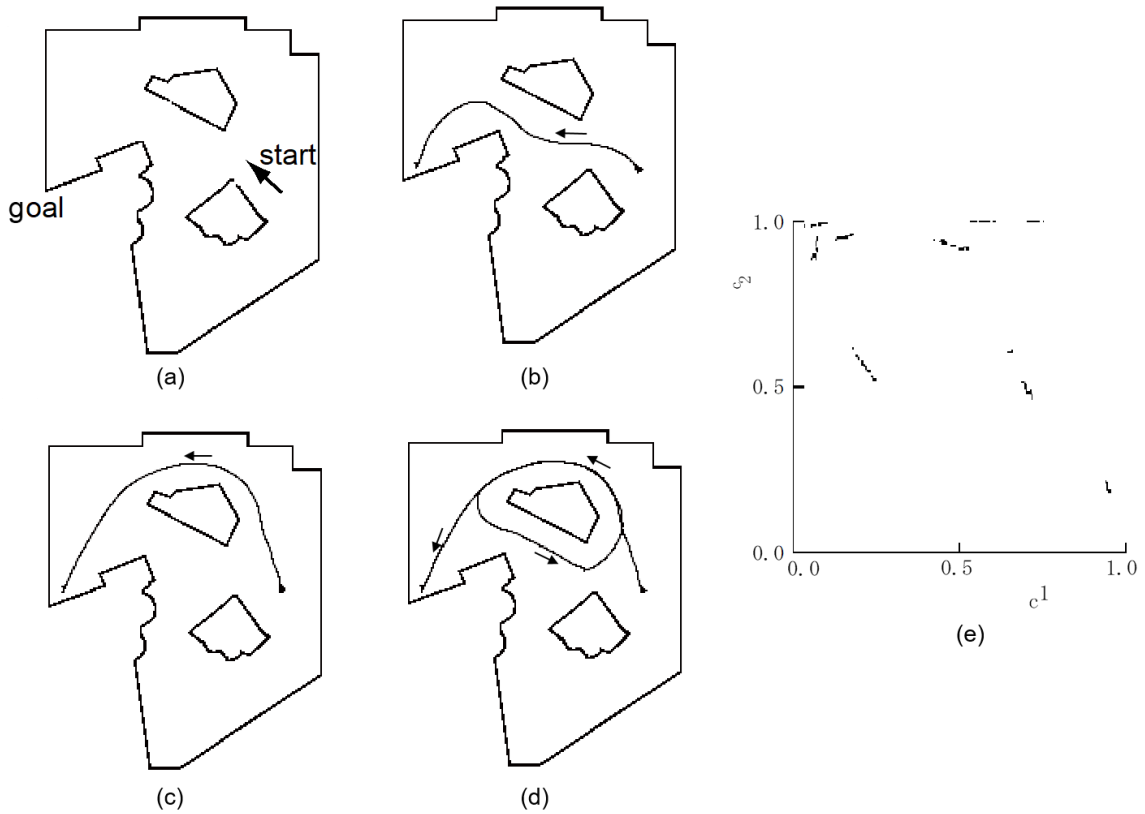681   segments as nodes of a state transition diagram.

682

Figure 5: Trajectories generated by goal-directed planning and the phase plot. (a) shows
staring and goal positions (b-d) 3 different trajectories generated by Yamabico based on
different action plans. (e) The phase plot of the internal state projected in a two-
dimensional space.

688

23

689   Importantly, each node is represented not by a point, but by an assembly of points in the
690   continuous state space of the RNN, because each point within a segment can have a
691   different history of branching in arriving there. If two points are neighboring each other,
692   two branching sequences of reaching these two points may be exactly the same for a
693   long past history. On the other hand, if two points are distant in the same segment, two
694   branching sequences reaching these points will be quite different in the immediate past
695   history. Since an infinite number of different branching sequences can be composed to
696   reach a branching point, and all points corresponding to those different compositions
697   should be embedded in a segment of finite length, Cantor set-like assembly is organized
698   within each segment. Cantor sets are interesting in this context because they represent
699   boundary points. They are perfect sets, meaning that a set is equal to its derivative set,
700   which means that it is equal to its limit points. These findings are analogous to what
701   Elman (1991) and Pollack (1991) showed in investigating the capability of RNNs to
702   learn word or symbol sequences regulated by grammar.

703

704   Further analysis revealed that the whole assembly of points in the phase plot represents
705   an invariant set of a global attractor. When mental simulation is perturbed with external
706   noise added to the network activity, prediction goes wrong and the internal state falls
707   out of the invariant set.[5] However, if the noise is removed, the predictability can be
708   recovered after several steps of mental simulation, and the trajectory of the internal state
709   returns back to the invariant set. In this way, the invariant set may represent the
710   boundary of cognition (Maturana and Varela, 1991) which is structurally stable against
711   perturbation due to the nature of an attractor.

712

713   One drawback in this study (Tani, 1996) is that the exploration of the environment was
714   conducted randomly, independent of the process of learning about the environment in
715   order to satisfy a purpose. Human infants or artificial agents may explore the world by
716   seeking some intrinsic rewards such as novelty in interacting with their environments
717   (Schmidhuber, 1991; Oudeyer et al., 2007; Tschantz et al., 2020). Tani and Yamamoto
718   (2002) extended the study in Tani (1996) in order to investigate the issue of intrinsic
719   motivation by adding a novelty rewarding mechanism as a drive for seeking novel
720   experiences during exploration. The exploration and the learning phases of this
721   experiment were interleaved, with each happening after the other. In the exploration
722   phase, action plans for branching sequences were generated and executed such that the

---

[5] The network loses its "grip" cf. Bruineberg et al., 2014.

sum of expected prediction error at each branching point was maximized (rather than minimized). And in the learning phase, the training of the network was conducted by using two types of sensory-action sequence data. One was the sequence which had been experienced in the last travel and which had been stored in the short-term memory, and the other was the set of rehearsed sequences generated by using the same planning scheme for maximizing the novelty (in which the sum of expected prediction error at each branching point was maximized). By extending the model thusly, Yamamoto and Tani showed that compositionality - in terms of the number of different combinations of branching - increases both in the physically generated trajectories and in the rehearsed sequences during the learning phase over the course of development. Finally, a phase space analysis of the internal state at each stage of development indicated that a segment-wise invariant set similar to the one shown in Figure 5 (e) appeared, but only during the end period of the development, when the robot had completely learned about all possible branching consequences. Such dynamics illustrate how the aforementioned boundary of cognition might emerge during developmental processes similarly motivated by exploration in biological models including human beings, as well.

Recently, Friston and colleagues (Kaplan and Friston, 2018) took a generic Bayesian perspective on the exploration-exploitation trade-off by considering that the expected free energy is comprised of novelty, salience and prior preferences which together constitute optimal beliefs on policies guiding action. On this scheme, novelty and salience represent intrinsic values where novelty is for epistemic gain, as described previously, saliency is for the gain of certainty about the latent state, and prior preferences serve the role of extrinsic values, such as for achievement of predefined rewards or goals. As such, Kaplan and Friston (2018) assumes that actions are selected both in order to gain knowledge about the world and to achieve predefined preferences. One point of interest in this context, however, is the relatively high computational cost required in order to optimize such action plans. In reality, it is not necessary to always compute optimal action plans, since action can typically be generated according learned routine, or habit. On this point, Maisto et al. (2019) proposed that an active inference agent caches the probabilities of policies from previous trials in memory as habits in order to reduce the computational costs for re-calculating them at each new trial. Simply put, a probability is only re-calculated when encountering a new context, and is then kept cached as long as the context does not change. Testing this sort of idea in robots should be practically beneficial in for example reducing real-time computational burdens of robots in operational contexts.

3. The sense of self

Building from the preceding introduction, this section speculates on how such cognitive robotics experiments employing fundamental principles may contribute to ongoing inquiry into the sense of self and related psychological phenomena. Although the problem of self has been addressed in various ways, such as Neisser's (1988) five different types of selves or Kohut's (2013) bipolar self from psychoanalysis, the current paper focuses on two distinct types of self, minimal self and narrative self. This section begins with Gallagher's (2000) concept of the minimal self, and then examines how his notions may corresponded to the phenomena observed in CNR experiments such as those introduced in the preceding section.

3.1 The minimal self

Gallagher (2000) argues that after all of the unessential features of experience are stripped away, we still have a feeling of a basic, immediate, or primitive 'something' that we can call the "minimal self". He further contemplates that this sort of non-reflective self is associated with two different types of senses, one is a sense of ownership and the other is a sense of agency. According to Gallagher (2000), the sense of ownership is the sense that I am the one who is undergoing an experience. For example, a sense that this is my body moving regardless of whether the movement is caused by me or others. The sense of agency, on the other hand, refers to congruence between an agent's intention or belief in an action and its anticipated outcome, which endows the agent with the sense that "I am the one generating this action".

Both cases may be explained in terms of internal models for predicting perceptual outcomes. For example, Hohwy (2013) showed that the sense of body ownership in the rubber hand illusion (Botvinick and Cohen, 1998) can be explained by using predictive coding that models the probabilities of the dummy hand being mine or another's. During the rubber hand illusion, experience of the temporally synchronized multimodal sensation of touching, one from a tactile stimulus and the other from visual observation, results in a causal inference that these two occur at the same location even though they actually do not. This is because the temporal correlation entails more dominant effects on the inference than the spatial one does (Hohwy, 2013: Limanowski and Blankenburg, 2013). This leads to the illusion of the dummy hand being mine.

794 Concerning the sense of agency, Gallagher (2000) suggests that a possible underlying
795 mechanism may be conceivable by considering neurocognitive models accounting for
796 some cases of schizophrenia such as discussed by Feinberg (1978) and Frith (1992). They
797 proposed that the delusion of control as a characteristic of schizophrenia may occur when
798 a mismatch takes place between an intended state and the anticipated state produced by
799 the forward model. As the forward model is informed by the motor efference copy, they
800 proposed that the mismatch may be caused by either the failure of the forward model, or
801 due to the fact that the efference copy cannot be sent to the forward model because the
802 motor controller is disconnected from it. Gallagher (2000) suggested that the sense of
803 agency, which remains implicit in a normal condition, can be disturbed in such a case,
804 resulting in the feeling that 'somebody is controlling me' which is common to self-reports
805 in some cases of schizophrenia.

806

807 The RNNPB robotics experiments (reviewed in section 2) may help to account for the
808 emergence of minimal self accordingly. Minimal self should develop implicitly as an
809 aspect of the sense of agency derived from causality between an agent's intention driving
810 action (encoded in the PB) and the affected perceptual reality, similarly to the sentiment
811 expressed in Hohwy (2007, page 5): "mineness is the feeling of already being familiar
812 with the movement's sensory consequences when they actually occur, we are so to speak
813 already 'at home' in the movement because the incoming signals are predicted through
814 habituation, and therefore it is regarded as an implicit sense of self." Accordingly, when
815 a robot's prediction was accurate, action proceeded smoothly and automatically without
816 distinction between synchronized embodied self and the external objects with which it
817 was interacting, such as a ball or human counterpart.

818

819 Our proposal in the context of minimal self is that, when such synchrony breaks down
820 due to miscellaneous unpredictable influences including noise in the physical system or
821 a human participant's sudden intentional change, the otherwise implicit sense of minimal
822 self becomes an object of consciousness. This is because the consequent effort required
823 to minimize the reconstruction error in the immediate past window, by inferring an
824 optimal intention state in the PB, is accompanied by a focal awareness of the gap between
825 embodied routines and the capacity for embodied routines to successfully meet
826 environmental demands. At this very moment of the unified structure breaking down, the
827 independence of each element becomes noticeable. Consequently, this experience of self

828   is formally articulated as *minimal self-consciousness* (Tani, 1998; 2016, pages 169-172).

829

830   Moreover, on this account, the sense of minimal self should intermittently shift between

831   an unconscious phase (predictable phase, when intention guides action without undue

832   error) and a conscious phase (unpredictable phase, when error forces reformulation of

833   intentions guiding action going forward) as had been observed in a vision-based robot in

834   Tani (1998), as well as in the humanoid robot experiments involving ball handling

835   described in section 2. Why do those system dynamics once converged into an attractor

836   basin, such as a predictable or routine interaction in a region, get destabilized again and

837   move out to another basin of attractor? One possibility is the inherently indeterministic

838   nature of embodied cognitive systems due to the circular causality established in the

839   enactment loop (Tani, 1998; 2009).

840

841   Circular causality describes the embedded and embodied agent's situation. An agent acts

842   on the world, and a sequence of causes and effects returns back to the original cause with

843   possibly altering it whereby another sequence is produced in on ongoing interactive

844   feedback loop. For example, in the case of the ball handling humanoid robot, when a

845   prediction error is generated for the ball position, the intention of the robot in terms of the

846   PB value is updated in the direction of minimizing the error. This intention generates the

847   next step prediction of proprioception which, in the case of tracking a ball, turn out to be

848   new target joint angles. These angles are fed into the robot motor controllers. Then, both

849   hands of the robot move to push the ball further, for example.

850

851   When every process in this loop proceeds ideally, the whole system dynamics stay always

852   in the same attractor basin by successfully minimizing prediction error, unless

853   exceptionally large noise comes from the external world. However, in reality under

854   resource bounded situations, this cannot be guaranteed. Inherent indeterminism will

855   appear. The predictability of the neural network is limited, because it is trained with only

856   a finite amount of the sensory-motor experience. The inference of the PB values cannot

857   be guaranteed to be optimal in a real time situation, and physical movements of the robot

858   body as well as the resultant movement of the ball should contain some margin of

859   unpredictability because of nonlinearity in both static and dynamic friction and contact

860   dynamics. All of these contingencies (to a very large degree) are due to embodiment.

861   They provide potential instability to the system dynamics.

862

863   In fact, two flows in opposing directions coexist. One is the inflow converging to
864 attractors or thermodynamic equilibria by minimizing prediction error (or free energy),
865 and the other is the outflow destabilizing the convergence by means of embodiment and
866 circular causality. In a macroscopic sense, the coexistence of inflow and outflow in the
867 phase space makes attractors only marginally stable, wherein the state trajectories tend to
868 visit multiple pseudoattractors one by one itinerantly (known as chaotic itinerancy, Tsuda
869 et al., 1987; Kaneko, 1990). This may correspond to the "momentary self" contemplated
870 by William James (1890) who wrote that: "When we take a general view of the wonderful
871 stream of our consciousness, what strikes is the pace of its parts. Like a bird's life, it
872 seems to be an alternation of flights and perchings (p. 243)."[6]

873

874   Such unsteady dynamics resulting from potential indeterminism provides an inherent
875 autonomy to the minimal self in terms of its spontaneous shifts between its unconscious
876 phase (staying inside basins of attractors or habitual regions) and conscious phase
877 (transition to another attractor passing through less familiar regions). Froese and Taguchi
878 (2019) present an analogous argument that artificial as well as living agents may make
879 sense of their interactions with the world provided that there is some room for
880 indeterminism or incompleteness in the causal closure of these interactions. One
881 difference, however, between Froese and Taguchi's (2019) and our consideration is that
882 they attribute the origin of indeterminacy to quantum mechanics at the micro level, which
883 is supposed to be amplified through the enactment loop at the macroscopic level. On the
884 contrary, we presume that the sensitivity to the initial state caused by chaos, or the
885 structural instability observed in chaotic itinerancy, may account for the origin of
886 indeterminacy without resorting to (what currently remain) mystic propositions.

887

888 **3.2 Narrative self**

889 Gallagher (2000) considers narrative self as "a more or less coherent self (or self-image)
890 that is constituted with a past and a future in the various stories that we and others tell

---

[6] In terms of compositionality, this image also recalls Plato's aviary from the Thaetatus, in which a
person collects birds in the cage of the mind (where they may be imagined to flit from perch to perch),
representing expressions of knowledge that can be taken when the knowledge is useful, or mistaken when
the knowledge is not, cf. discussion beginning 197d, Plato, 1997, page 218.

891   about ourselves" (p.15). He contrasts two distinct ways of representing such a sense of

892   narrative self. One is offered by Dennett (1992) who characterizes self as the constant

893   locus of experience and center of "narrative gravity". The other is a more distributed

894   model inspired by Paul Ricoeur's (1984) philosophy of narrative self. Ricoeur considers

895   a hermeneutic cycle of movements from *prefiguration* of phenomena in the world to their

896   *refiguration* or restoration back into the real world through (communicative) action via

897   *configuration* of interpreted narratives, with configuration of narratives playing the role

898   of mediation between prefiguration and refiguration, and the three together constituting a

899   process through which the agent gains a better understanding of its self and its place in

900   the social and natural world. Especially, he considers the human experience of aporia,

901   when phenomena in the world are experienced as incomprehensibly contradictory.

902   Ricoeur emphasizes that human beings compose fictive as well as true narratives in order

903   to reconcile this feeling of aporia. Furthermore, he considers that one's own self-

904   narratives are configured in a way that they are intermingled with those communicated

905   by others. Following Ricoeur, Gallagher (2000) proposes that narrative self might be

906   developed as a mixing of stories about one's self, including conflictive and irresolvable

907   ones which an individual might tell about her/himself or others might tell about her/him

908   (cf. "pernicious misunderstandings" in Gallagher and Allen, 2018, pages 14-15) Thus on

909   Gallagher's account, the sense of narrative self might be considered as a center of

910   narrative gravity that is more distributed, representing the reconciliation of narratives

911   normalized in communication with others in interaction with the shared object

912   environment.

913   We can find some analogy with what Ricoeur and Gallagher consider narrative,

914   including this more distributed sense of narrative self, in the results of the CNR

915   experiments described previously. In effect, the robots used in experiments by Tani and

916   colleagues were frequently confronted with incomprehensible and irresolvable

917   situations during tutoring. For example, the humanoid robot implemented with the

918   MTRNN (see section 2.3) was tutored to generate inconsistent primitive movement

919   sequences, e.g. after grasping an object, it was tutored to lift the object up in one

920   instance, and in another it was tutored to rotate the same object after grasping it. The

921   mobile robot Yamabico was tutored to branch to the left during one trial, and during

922   another to branch right at the same branching point. Each tutoring trajectory never

923 repeated exactly the same sensory-motor sequence pattern as another, even though the
924 tutors attempted to do so, due to noise and fluctuation – indeterminacy (cf. Tani, 1998;
925 Froese and Taguchi, 2019) – inevitable during embodied interaction with the physical
926 world. In light of Gallagher and Ricoeur's insights into narrative self, we may say that
927 these robots could develop compositionality for generating diverse actions because it
928 was necessary to deal with the inconsistency presented in the tutoring trajectories.
929 Through consolidative learning, the RNNs could self-organize finite state machine-like
930 state transition structures in their latent state space by extracting relational structures
931 from among the tutored set of inconsistent sequences.
932
933 The segmentation and chunking of the continuous sensory flow observed in robotic
934 experiments using the RNNPB and MTRNN are mechanized by means of the error
935 regression employed in order to minimize the error signal generated in prediction and
936 reconstruction at the moment of transition from one primitive to another. The
937 inconsistency brought to the robots especially by human tutors in demonstration and
938 tutoring of movement patterns plays an important role here, also. If we tutor the robots
939 with a continuous sensory flow consisting of a primitive sequence like $A \rightarrow B \rightarrow C$,
940 repeated without inconsistencies, this sensory flow will be learned as a big chunk
941 without segmentation into primitives. But, if the robots are tutored with a set of sensory
942 flows consisting of inconsistent primitive sequences like $A \rightarrow B \rightarrow C$, $A \rightarrow B \rightarrow A$, and
943 $A \rightarrow C \rightarrow B$, these sensory flows are learned as compositions of reusable primitives with
944 each of them segmented. We presume that the cognitive competency of
945 compositionality for segmenting the continuous sensory flow into a set of reusable
946 primitives and composing and decomposing a whole action sequence using these
947 primitives can provide a basis for development of the narrative self.
948
949 Furthermore, some robots (or more generally agents) (Schmidhuber, 1991; Tani and
950 Yamamoto, 2002; Oudeyer 2007; Kaplan and Friston, 2018) are motivated to learn to
951 predict unpredictable situations by seeking novelty, with the objective of which also
952 seeming rather contradictory or conflictive at first glance. And, in order to reconcile
953 such conflicts, agents seem to be required to generate creative or fictive mental images
954 accounting for the hidden causal structure in the world. Here, we might note some
955 analogy between the compositionality developed by means of a self-organizing finite
956 state machine-like structure in distributed neural activation in the RNN during such
957 conflictive situations, and the sense of narrative self which develops distributedly, with
958 the mixing of diverse inputs including inconsistent ones from the outside as described

31

959   above. The compositionality developed in the robots enables them to mentally simulate
960   future actions, including counter factual or fictional ones, as well as to rehearse what
961   has been experienced in the past in order to prepare for the uncertain future.
962
963   At this point, it is natural to consider that the development of narrative requires
964   linguistic competency for telling stories. On this front, there have been some efforts in
965   attempting to ground linguistic expressions in sensory-motor modalities by using
966   RNNPB and MTRNN architectures by some groups (Sugita and Tani, 2005; Peniak et
967   al., 2011; Yamada et al., 2016; Heinrich and Wermter, 2018). With a vision-based
968   mobile robot using the predictive coding framework, Sugita and Tani (2005) showed
969   that an RNNPB can learn to bind a set of simple imperative sentences consisting of
970   verbs and nouns, e.g. Point-Red, Push-Blue, Hit-Green, with corresponding sensory-
971   motor behavioral patterns. An analysis on the experimental results of learning and
972   action generation for given imperative sentences showed that the compositionality in
973   combining verbs and nouns in the linguistic modality and the one in combining actions
974   and objects developed as a unified structure in the RNNPB. Peniak et al. (2011) and
975   Heinrich and Wermter (2018) showed scaling of such language-behavior binding using
976   extended MTRNN models. Also extending the MTRNN, Yamada et al. (2016)
977   presented a continuing human-robot interaction experiment using both linguistic and
978   behavioral modalities. These experimental results revealed that the contextual flow
979   corresponding to successive human-robot interactions was represented in the higher-
980   level latent variables in the MTRNN. These results suggest that the internal structures
981   developed in the latent space in these RNN models, via continuing human-robot
982   interaction using both linguistic and behavioral modalities, bring us *closer* to realizing a
983   narrative self as articulated by Ricoeur and by Gallagher and colleagues in an embodied
984   cognitive neurorobot.
985
986   **4. Discussion**
987
988   The preceding paper introduced cognitive neurorobotics, the principles of prediction
989   error minimization and backpropagation as implemented in different RNN architectures,
990   related these with free energy and active inference, and reviewed selected cognitive
991   neurorobotics experiments employing these principles in greater detail. In terms of
992   Marr's (1982) three levels introduced in section 1, the error minimization principle
993   appears at the computational level, prediction and active inference at the algorithm
994   level, and RNNs embodied by robots at the implementation level as an example. This

995 leaves open questions, for instance how these implementations can be validated, and
996 then to what extent we may expect them to contribute to scientific explanation of
997 cognitive phenomena including the senses of minimal and narrative self as reported by
998 human beings.
999

1000 First, in the review of the CNR experiments using the RNNPB (Ito and Tani, 2004; Ito
1001 et al., 2006), we explained that the RNNPB learns multiple behaviors in the course of
1002 prediction error minimization as embedded in different attractor basins which represent
1003 habitual regions for the robot. It was also explained that behavior patterns of the robots
1004 shift from a learned one, by means of the error regression accompanied by segmentation
1005 of continuous sensory-motor flows. to another due to either external forces or internal
1006 fluctuation
1007

1008 Next, in the review of the CNR experiments using the MTRNN (Yamashita and Tani,
1009 2008; Nishimoto and Tani, 2009), it was explained that compositionality as a cognitive
1010 competency for composing/decomposing the whole action from/into behavior
1011 primitives can develop gradually in the course of iterative tutoring of the robot by using
1012 the error minimization principle. We explained that such compositionality can develop
1013 by means of self-organization of functional hierarchy using the prior constraints applied
1014 to the network, including layer-wise timescale difference and the information
1015 bottlenecks in the connectivity between layers.
1016

1017 In the review of experiments on goal-directed planning in robot navigation tasks, it was
1018 explained that robots can learn compositional structures latent in the obstacle
1019 environment through either supervised tutoring (Tani, 1996) or self-exploration (Tani
1020 and Yamamoto, 2002). Further, it was noted that such compositional structures develop
1021 by self-organizing global attractors of Cantor set-like assembly in the latent state phase
1022 space. Although its appearance seems analogous to finite state machines at first glance,
1023 these two are crucially different. As for the attractors developed in the latent state phase
1024 space, they represent the boundary of cognition (Maturana and Varela, 1991) wherein
1025 prediction goes well as habituated within the invariant set of the attractors, and
1026 prediction goes wrong once the state trajectory goes out of the invariant by possible
1027 permutation. However, the state trajectory can come back to the invariant set as long as
1028 it is formed as a global attractor. On the other hand, in the case of a finite state machine,
1029 there is no mechanism for such auto-recovery unless some external programs for this
1030 purpose are provided. The CNR experiments reviewed above demonstrate that robots

1031 can generate both fictional and factual action plans by using compositional structures
1032 developed in both goal-directed planning and novelty rewarding schemes.
1033
1034 Further analysis contemplated possible accounts for subject experience including senses
1035 of minimal and narrative self. First, let us revisit our account of the sense of minimal
1036 self. When action goes smoothly and an agent remains in habitual regions by
1037 minimizing the prediction error, the sense of minimal self is present but only implicitly.
1038 However, a breakdown of such a steady phase comes inevitably because of the inherent
1039 indeterminacy due to the circular causality established in the enactment loop. In such an
1040 instance, the minimal self should become an object of conscious awareness with the
1041 effort to return from unfamiliar regions to a routine one by minimizing prediction error.
1042 We see the structure of the minimal self in this autonomy of spontaneous shifts between
1043 unconscious and conscious phases analogous with James' (1890) "wonderful stream of
1044 our consciousness".
1045
1046 Next, let us revisit the sense of narrative self. We found a good analogy in the results of
1047 the CNR experiments with what Ricoeur and Gallagher's socially distributed sense of
1048 narrative self. Robotic experiments using the RNNPB and MTRNN showed that
1049 compositionality can be naturally developed provided that the robots are tutored with a
1050 set of inconsistent sensory-motor sequences, corresponding with Ricoeur's thought that
1051 humans compose both fictive and true narratives in the process of resolution of
1052 aporia. Indeed, the CNR experiments reviewed in the current paper showed that these
1053 robots can generate both fictive and factual compositions of primitive action both in
1054 physical execution and in mental planning and rehearsing. And, by briefly introducing
1055 the on-going research on embodied language using various RNN models, discussion
1056 extended to possibilities of how such narratives initially represented in distributed
1057 neural activation patterns can be transformed into linguistic representations for sharing
1058 stories.
1059
1060 Although space forbids present review, other ongoing work "breaks" these and
1061 complimentary architectures (cf. Glennan, 2005), tests them against biological models
1062 in similar abnormal conditions and in this way aims to inform accounts of psychiatric
1063 conditions e.g. schizophrenia and autism understood as self-disturbances (Yamashita
1064 and Tani, 2012; Idei et al., 2018). It is expected that such research contributes to our
1065 understanding of otherwise difficult to resolve cognitive phenomena in two directions.
1066 For one, emergent dynamics when analogous with biological model behavior may

34

1067 inform researchers working at lower levels of organization about how different
1068 operations may be related, e.g. temporally. In such an instance, prediction error
1069 minimization inspires neural network algorithm design (backpropagation, error
1070 regression, etc.). With successful demonstration of functional dynamics in a real-world
1071 context, such a synthetic architecture can be correlated with biological models (e.g., the
1072 higher, intermediate, and the lower level correspond to the PFC, parietal, and
1073 S1+V1+M1) in exploration of possible explanations for (normal and abnormal)
1074 biological structural dynamics. An example of such work includes the HBP's
1075 neurorobotics group's effort to render a computational model of an embodied mouse
1076 that may replace biological models in psychological studies (Falotico et al., 2017).
1077
1078 On the other hand, the neurorobotics studies reviewed in this paper target invariant
1079 structures arising in what Sun et al. (2005) call the "causal nexus" between top-down
1080 and bottom-up processes, e.g. selves, internal world-models. Informed by
1081 phenomenological and neurocognitive research, these studies aim to contribute to
1082 explanations in cognitive science by articulating architectures which generate target
1083 emergent phenomena through their dynamic interaction with the world. In such an
1084 experiment for example, the prediction error minimization principle shapes the
1085 cognitive architecture according to biological and psychological constraints. This
1086 architecture is tested in robot experiments, and emergent phenomena are recorded.
1087 Correspondences to phenomenology (e.g., minimal-self), to neuroscience (e.g. mirror
1088 neurons), and to psychiatry (e.g. autism) are considered. Limitations of the model
1089 inform ongoing inquiry, and insights drawn from these studies are offered.
1090
1091 Although physical robot experiments might be replaced by simulations in some
1092 contexts, it must be stressed that embodied humanoid robots are important to the
1093 success of the CNR research such as that reviewed in section 2. Practically, the results
1094 of physical robot experiments are more robust than simulated variants. Because
1095 simulation experiments typically employ thousands of trials that establish optimal
1096 parameters within narrow ranges, the resulting networks often become too rigid to
1097 perform in real-world embodied robots because robots themselves are quite noisy (in
1098 the informational sense) in their mechanics and physical interactions. With this in mind,
1099 we may point to an epistemic upshot to embodied robotics experiments over simulated
1100 variants. Humanoid CNR experiments open phenomena to investigation which remain
1101 inaccessible to simulations, especially those which emerge through direct human
1102 interaction as in the experiments. During these experiments, trajectories of every

1103 learned movement are co-developed by tutor and robot through their interaction.
1104 Accordingly, we speculate that successful results were achieved not only due to
1105 essential structural dynamics captured by biologically inspired cognitive architectures,
1106 but also due to the intuitive interaction afforded human tutors with both tutor and
1107 humanoid robot aiming to minimize error in the embodied manipulation of common
1108 objects – including each other - in a shared space of action. Such phenomena cannot be
1109 (easily) simulated. And, given the fundamental role of interaction with others in the
1110 development of self (cf. Bolis and Schilbach, 2020), the socially situated nature of
1111 embodied humanoid cognitive neurorobotics experiments presents special potential for
1112 ongoing inquiry into phenomena associated with self in human beings.
1113
1114 **4. Conclusion**
1115
1116 Guided by the intuition that higher-level cognitive phenomena including different
1117 senses of self should emerge from the effort to minimize conflicting interactions
1118 between top-down and bottom-up information processes, Tani and colleagues have been
1119 refining cognitive neurodynamic models since the mid-1990s to articulate structural
1120 dynamics native to living systems in artificial ones. The preceding paper reviewed
1121 synthetic neurorobotics experiments using analog devices (RNNs) directly sharing the
1122 same analog metric space with human beings with the expectation that such continuous
1123 spatio-temporal dynamics can both avoid the notorious symbol grounding problem
1124 (Harnad, 1990) as well as inform our understanding of cognitive phenomena such as
1125 self in human beings.
1126
1127 **Works consulted:**
1128

1129 Ahmadi, A., & Tani, J. (2019). A Novel Predictive-Coding-Inspired Variational RNN
1130 Model for Online Prediction and Recognition. Neural Computation, 31, 2025–2074.
1131
1132 Arbib, M. A. (1981). Perceptual structures and distributed motor control. In V. B.
1133 Brooks (Ed.), Handbook of physiology, the nervous system II. Motor control (pp. 1448-
1134 1480). Bethesda, MD: American Physiological Society.
1135
1136 Arie, H., Endo, T., Arakaki, T., Sugano, S., & Tani, J. (2009). Creating novel goal-
1137 directed actions at criticality: a neuro-robotic experiment. New Mathematics and
1138 Natural Computation, 5(1), 307-334.

1139

1140 Asada, M., Hosoda, K., Kuniyoshi, Y., Ishiguro, H., Inui, T., Yoshikawa, Y., ... and

1141 Yoshida, C. (2009). Cognitive developmental robotics: A survey. IEEE transactions on

1142 autonomous mental development, 1(1), 12-34.

1143

1144 Baltieri, M., & Buckley, C. L. (2017). An active inference implementation of

1145 phototaxis. In Artificial Life Conference Proceedings 14 (pp. 36-43). One Rogers Street,

1146 Cambridge, MA 02142-1209 USA journals-info@ mit. edu: MIT Press.

1147

1148 Baltieri, M., & Buckley, C. L. (2018). A probabilistic interpretation of PID controllers

1149 using active inference. In International Conference on Simulation of Adaptive Behavior

1150 (pp. 15-26). Springer, Cham.

1151

1152 Barandiaran, X. E., & Chemero, A. (2009). Animats in the Modeling

1153 Ecosystem. Adaptive Behavior, 17(4), 287-292.

1154

1155 Beer, R. D. (2000). Dynamical approaches to cognitive science. Trends in cognitive

1156 sciences, 4(3), 91-99.

1157

1158 Bolis, D., & Schilbach, L., (2020). 'I Interact Therefore I Am': The Self as a Historical

1159 Product of Dialectical Attunement. Topoi, 39(3), 521-534.

1160

1161 Botvinick, M., & Cohen, J. (1998). Rubber hands 'feel' touch that eyes see. Nature,

1162 391(6669), 756-756.

1163

1164 Brooks, R. A. (1990). Elephants don't play chess. Robotics and Autonomous

1165 Systems, 6, 3– 15.

1166

1167 Brooks, R. A. (1991). Intelligence without representation. Artificial Intelligence

1168 Journal, 47, 139– 159.

1169

1170 Bruineberg, J., Rietveld, E., Bruineberg, J., Rietveld, E., Bruineberg, J., & Rietveld, E.

1171 (2014). Self-organization, free energy minimization, and optimal grip on a field of

1172 affordances. Frontiers in Human Neuroscience, 8.

1173

1174 Bruineberg, J., Rietveld, E., Parr, T., van, M. L., & Friston, K. J. (2018). Free-energy
1175 minimization in joint agent-environment systems: A niche construction
1176 perspective. Journal of Theoretical Biology, 455, 161-178.
1177
1178 Cangelosi, A., & Schlesinger, M. (2015). Developmental robotics: From babies to
1179 robots. MIT press.
1180
1181 Chao, Z. C., Takaura, K., Wang, L., Fujii, N., & Dehaene, S. (2018). Large-scale
1182 cortical networks for hierarchical prediction and prediction error in the primate brain.
1183 Neuron, 100(5), 1252-1266.
1184
1185 Clark, A. (2015). Radical Predictive Processing. The Southern Journal of
1186 Philosophy, 53, 3-27.
1187
1188 Costello, F., & Keane, M. (2000). Efficient creativity: Constraint-guided conceptual
1189 combination. Cognitive Science, 24 (2), 299–349.
1190
1191 Dennett, Daniel C. (1992) The Self as a Center of Narrative Gravity. In: F. Kessel, P.
1192 Cole and D. Johnson (eds.) Self and Consciousness: Multiple Perspectives. Hillsdale,
1193 NJ: Erlbaum.
1194
1195 Doya, K., Samejima, K., Katagiri, K. I., & Kawato, M. (2002). Multiple model-based
1196 reinforcement learning. Neural computation, 14(6), 1347-1369.
1197
1198 Elman, J. L. (1991). Distributed representations, simple recurrent networks, and
1199 grammatical structure. Machine learning, 7(2-3), 195-225.
1200
1201 Evans, G. (1982). The varieties of reference. Oxford: Clarendon Press.
1202
1203 Falotico, E., Vannucci, L., Ambrosano, A., Albanese, U., Ulbrich, S., Vasquez, T. J. C.,
1204 Hinkel, G., ... Gewaltig, M. O. (2017). Connecting Artificial Brains to Robots in a
1205 Comprehensive Simulation Framework: The Neurorobotics Platform. Frontiers in
1206 Neurorobotics, 11.
1207
1208 Friston, K. (2005). A Theory of Cortical Responses. Philosophical Transactions:
1209 Biological Sciences, 360(1456), 815-836.

1210

1211 Friston, K., Mattout, J., and Kilner, J. (2011). Action understanding and active

1212 inference. Biological cybernetics, 104 (1-2), 137-160.

1213

1214 Frith, C. D. (1992). The cognitive neuropsychology of schizophrenia. Psychology press.

1215

1216 Feinberg, I. (1978). Efference copy and corollary discharge: implications for thinking

1217 and its disorders. Schizophrenia bulletin, 4(4), 636.

1218

1219 Froese, T., & Ziemke, T. (2009). Enactive artificial intelligence: Investigating the

1220 systemic organization of life and mind. Artificial Intelligence, 173(3-4), 466-500.

1221

1222 Froese, T., & Taguchi, S. (2019). The Problem of Meaning in AI and Robotics: Still

1223 with Us after All These Years. Philosophies, 4 (2), 1-14.

1224

1225 Gallagher, S. (2000). Philosophical conceptions of the self: implications for cognitive

1226 science. Trends in Cognitive Sciences, 4, (1), 14-21.

1227

1228 Gallagher, S., & Allen, M. (2018). Active inference, enactivism and the hermeneutics of

1229 social cognition. Synthese, 195(6), 2627-2648.

1230

1231 Geldard, F. A., & Sherrick, C. E. (1972). The cutaneous "rabbit": a perceptual illusion.

1232 Science, 178(4057), 178-9.

1233

1234 Glennan, S. (2005). Modeling Mechanisms. Studies in History and Philosophy of

1235 Science Part C: Studies in History and Philosophy of Biological and Biomedical

1236 Sciences, 36, 443-464.

1237

1238 Harnad, S. (1990). The symbol grounding problem. Physica D: Nonlinear Phenomena,

1239 42(1-3), 335-346.

1240

1241 Heinrich, S., & Wermter, S. (2018). Interactive natural language acquisition in a multi-

1242 modal recurrent neural architecture. Connection Science, 30(1), 99-133.

1243

1244 Hohwy, J. (2007). The sense of self in the phenomenology of agency and perception.

1245 Psyche, 13(1), 1-20.

1246

1247   Hohwy, J. (2013). The predictive mind. Oxford University Press.

1248   James, W. (1890). The principles of psychology. Dover Publ. (reprinted 1950).

1249

1250   Holland, O. (2007). A strongly embodied approach to machine consciousness. Journal

1251   of Consciousness Studies, 14(7), 97-110.

1252

1253   Hwang, J., Kim, J., Ahmadi, A., Choi, M., & Tani, J. (2018). Dealing with large-scale

1254   spatio-temporal patterns in imitative interaction between a robot and a human by using

1255   the predictive coding framework. IEEE Transactions on Systems, Man, and

1256   Cybernetics: Systems, 99, 1-14.

1257

1258   Idei, H., Murata, S., Chen, Y., Yamashita, Y., Tani, J., & Ogata, T. (2018). A

1259   Neurorobotics Simulation of Autistic Behavior Induced by Unusual Sensory Precision.

1260   Computational Psychiatry, 2, 164-182.

1261

1262   Iizuka, H., & Di Paolo, E. A. (2007). Toward Spinozist robotics: Exploring the minimal

1263   dynamics of behavioral preference. Adaptive Behavior, 15(4), 359-376.

1264

1265   Ito, M., & Tani, J. (2004). On-line imitative interaction with a humanoid robot using a

1266   dynamic neural network model of a mirror system. Adaptive Behavior, 12(2), 93-115.

1267

1268   Ito, M., Noda, K., Hoshino, Y., & Tani, J. (2006). Dynamic and interactive generation

1269   of object handling behaviors by a small humanoid robot using a dynamic neural

1270   network model. Neural Networks, 19, 323-337.

1271

1272   James, W. (1890). The principles of psychology: In two volumes. New York: Holt.

1273

1274   Kaneko, K. (1990). Clustering, coding, switching, hierarchical ordering and

1275   control in a network of chaotic elements. Physica D, 41, 137– 72.

1276

1277   Kaplan, D. M. (2015). Moving parts: the natural alliance between dynamical and

1278   mechanistic modeling approaches. Biology and Philosophy, 30 (6), 757-786.

1279

1280   Kaplan, R., & Friston, K. J. (2018). Planning and navigation as active inference.

1281   Biological cybernetics, 112(4), 323-343.

1282

1283   Kawato, M. (1999). Internal models for motor control and trajectory planning. Current
1284   opinion in neurobiology, 9(6), 718-727.

1285

1286   Knoll, A. and Gewaltig, O. (2016) Neurorobotics: A strategic pillar of the Human Brain
1287   Project. in Brain inspired intelligent robotics: The intersection of robotics and
1288   neuroscience. Science/AAAS: Washington, DC, 25-34.

1289

1290   Kohut, H. (2013). The analysis of the self: A systematic approach to the psychoanalytic
1291   treatment of narcissistic personality disorders. University of Chicago Press.

1292

1293   Kotseruba, I., & Tsotsos, J.K. (2018) 40 years of cognitive architectures: core cognitive
1294   abilities and practical applications. Artificial Intelligence Review, 1-78.
1295   https://doi.org/10.1007/s10462-018-9646-y.

1296

1297   Krichmar J. L. (2018). Neurorobotics - A Thriving Community and a Promising
1298   Pathway Toward Intelligent Cognitive Robots. Frontiers in Neurorobotics, 12, 42.
1299   doi:10.3389/fnbot.2018.00042

1300

1301   Kuniyoshi, Y., & Sangawa, S. (2006). Early motor development from partially ordered
1302   neural-body dynamics: experiments with a cortico-spinal-musculo-skeletal model.
1303   Biological cybernetics, 95(6), 589.

1304

1305   Lang, C., Schillaci, G., and Hafner, V. V. (2018). A Deep Convolutional Neural
1306   Network Model for Sense of Agency and Object Permanence in Robots. In 8th Joint
1307   IEEE International Conference on Development and Learning and Epigenetic Robotics
1308   (ICDL-EpiRob). IEEE, 260-265.

1309

1310   Lanillos, P., Cheng, G., and 2018 IEEE/RSJ International Conference on Intelligent
1311   Robots and Systems, IROS 2018. (2018). Adaptive Robot Body Learning and
1312   Estimation Through Predictive Coding. IEEE International Conference on Intelligent
1313   Robots and Systems, 4083-4090.

1314

1315   Levesque, H., & Lakemeyer, G. (2008). Cognitive robotics. Foundations of artificial
1316   intelligence, 3, 869-886.

1317

1318   Levine, S., Finn, C., Darrell, T., & Abbeel, P. (2016). End-to-end training of deep
1319   visuomotor policies. The Journal of Machine Learning Research, 17(1), 1334-1373.
1320
1321   Limanowski, J., & Blankenburg, F. (2013). Minimal self-models and the free energy
1322   principle. Frontiers in human neuroscience, 7, 547.
1323
1324   Maisto, D., Friston, K., & Pezzulo, G. (2019). Caching mechanisms for habit formation
1325   in Active Inference. Neurocomputing, 359, 298-314.
1326
1327   Marr, D. (1982). Vision: A computational investigation into the human representation
1328   and proceedings of visual information. New York: Freeman.
1329
1330   Maturana, H. R., & Varela, F. J. (1991). Autopoiesis and cognition: The realization of
1331   the living (Vol. 42). Springer Science & Business Media.
1332
1333   Metta, G., Sandini, G., Vernon, D., Natale, L., & Nori, F. (2008). The iCub humanoid
1334   robot: an open platform for research in embodied cognition. In Proceedings of the 8th
1335   workshop on performance metrics for intelligent systems . ACM. 50-56.
1336
1337   Miall, R. C., & Wolpert, D. M. (1996). Forward models for physiological motor control.
1338   Neural networks, 9(8), 1265-1279.
1339
1340   Millan, J. D. R., Leeb, R., Rupp, R., Muller-Putz, G. R., Neuper, C., Murray-Smith, R.,
1341   Giugliemma, C., ... Kubler, A. (2010). Combining brain-computer interfaces and
1342   assistive technologies: State-of-the-art and challenges. Frontiers in Neuroscience, 4.
1343
1344   Miyazaki, M., Hirashima, M., & Nozaki, D. (2010). The "Cutaneous Rabbit" Hopping
1345   out of the Body. Journal of Neuroscience, 30, 5, 1856-1860.
1346
1347   Moxon, K. A., Foffani, G., & Foffani, G. (2015). Brain-machine interfaces beyond
1348   neuroprosthetics. Neuron, 86 (1), 55-67.
1349
1350   Murata, S., Yamashita, Y., Arie, H., Ogata, T., Sugano, S., & Tani, J. (2015). Learning
1351   to perceive the world as probabilistic or deterministic via interaction with others: a
1352   neuro-robotics experiment. IEEE Transactions on Neural Networks and Learning
1353   Systems, (4), 830-848.

1354

1355 Neisser, U. (1988) Five kinds of self-knowledge. Philos. Psychol. 1, 35–59

1356

1357 Nilsson, N. J. (1984). Shakey the robot. SRI INTERNATIONAL MENLO PARK CA.

1358

1359 Nishimoto, R., & Tani, J. (2009). Development of hierarchical structures for actions and
1360 motor imagery: a constructivist view from synthetic neuro-robotics study. Psychological
1361 Research, 73, 545-558.

1362

1363 Nolfi, S., Floreano, D., & Floreano, D. D. (2000). Evolutionary robotics: The biology,
1364 intelligence, and technology of self-organizing machines. MIT press.

1365

1366 Oudeyer, P. Y., Kaplan, F., & Hafner, V. V. (2007). Intrinsic motivation systems for
1367 autonomous mental development. IEEE transactions on evolutionary computation,
1368 11(2), 265-286.

1369

1370 Pastra, K., & Aloimonos, Y. (2012). The minimalist grammar of action. Philosophical
1371 Transactions of the Royal Society B: Biological Sciences, 367(1585), 103-117.

1372

1373 Peniak, M., Marocco, D., Tani, J., Yamashita, Y., Fischer, K., & Cangelosi, A. (2011).
1374 Multiple time scales recurrent neural network for complex action acquisition.
1375 Proceedings of ICDL-Epirob.

1376

1377 Piaget, J. (1953). The Origins of Intelligence (New York: Routledge).
1378 Vygotsky, L. (1962). Thought and Language (Cambridge MA: MIT Press) (original
1379 work published in 1934).

1380

1381 Plato, Cooper, J. M., & Hutchinson, D. S. (1997). Plato: Complete works. Indianapolis,
1382 Ind: Hackett.

1383

1384 Pollack, J. B. (1991). The induction of dynamical recognizers. Machine Learning, 7,
1385 227-252.

1386

1387 Prescott, T. J., & Camilleri, D. (2019). The synthetic psychology of the self. In
1388 Cognitive Architectures (pp. 85-104). Springer, Cham.

1389

1390 Raab, D. H. (1963). Backward masking. Psychological Bulletin, 60(2), 118

1391

1392 Rao, R., & Ballard, D. H. (1999). Predictive coding in the visual cortex: a functional
1393 interpretation of some extra-classical receptive-field effects. Nature Neuroscience, 2 (1),
1394 79-87.

1395

1396 Ricoeur, P. (1984) Time and Narrative (3 Vols), University of Chicago Press

1397

1398 Rumelhart, D. E., Hinton, G. E., Williams, R. J., & University of California, San Diego.
1399 (1985). Learning internal representations by error propagation. In Rumelhart, D. and
1400 McClelland, F., Parallel Distributed Processing, Vol. 1, Cambridge, MA: MIT Press.

1401

1402 Schillaci, G., Ritter, C. N., Hafner, V. V., & Lara, B. (2016). Body representations for
1403 robot ego-noise modelling and prediction. Towards the development of a sense of
1404 agency in artificial agents. In Proceedings of the Artificial Life Conference 2016.
1405 Cambridge, MA: MIT Press. 390-397.

1406

1407 Schmidhuber, J. (1991). Adaptive confidence and adaptive curiosity. In Institut fur
1408 Informatik, Technische Universitat Munchen, Arcisstr. 21, 800 Munchen 2.

1409

1410 Shimojo, S. (2014). Postdiction: its implications on visual awareness, hindsight, and
1411 sense of agency. Frontiers in psychology, 5(196).

1412

1413 Silberstein, Michael, & Chemero, Tony. (2013). Constraints on Localization and
1414 Decomposition as Explanatory Strategies in the Biological Sciences. Philosophy of
1415 Science, 80 (5), 958-970.

1416

1417 Sugita, Y., & Tani, J. (2005). Learning Semantic Combinatoriality from the Interaction
1418 between Linguistic and Behavioral Processes. Adaptive Behavior, 13 (1), 33-52.

1419

1420 Sun, R. (2002). Duality of the mind: A bottom-up approach toward cognition. London:
1421 Routledge

1422

1423 Sun, R., Coward, L. A., & Zenzen, M. J. (2005). On levels of cognitive
1424 modeling. Philosophical Psychology, 18 (5), 613-637.

1425

Tani, J. (1996). Model-based learning for mobile robot navigation from the dynamical systems perspective. IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), 26(3), 421-436.

Tani, J. (1998). An interpretation of the self from the dynamical systems perspective: a constructivist approach. Journal of Consciousness Studies, 5(5), 516-542.

Tani, J. (2003). Learning to generate articulated behavior through the bottom-up and the top-down interaction processes. Neural Networks, 16 (1), 11-23.

Tani, J. (2009). Autonomy of 'Self' at criticality: The perspective from synthetic neuro-robotics. Adaptive Behavior, 17(5), 421-443.

Tani, J., & Yamamoto, J. (2002). On the dynamics of robot exploration learning. Cognitive Systems Research, 3(3), 459-470.

Tani, J., Ito, M., and Sugita, Y. (2004). Self-organization of distributedly represented multiple behavior schemata in a mirror system: Reviews of robot experiments using RNNPB. Neural Networks, 17, 1273-1289.

Tani, J. (2016). Exploring robotic minds. Actions, symbols, and consciousness as self-organizing dynamic phenomena. Corby: Oxford University Press.

Tschantz, A., Seth, A., & Buckley, C. (2020). Learning action-oriented models through active inference. Plos Computational Biology, 16(4).

Tsuda, I., Körner, E. & Shimizu, H. (1987). Memory dynamics in asynchronous neural networks. Progress of Theoretical Physics, 78, 51– 71.

Vygotsky, L., Vakar, G., & Hanfmann, E. (1962). Thought and language. Cambridge, Mass: MIT Press.

Werbos, P. (1990). Backpropagation through time: What it does and how to do it. Proceedings of the IEEE, 78 (10), 1550-1560.

1461   Yamada, T., Murata, S., Arie, H., & Ogata, T. (2016). Dynamical integration of
1462   language and behavior in a recurrent neural network for human–robot interaction.
1463   Frontiers in neurorobotics, 10(5).
1464
1465   Yamashita, Y., & Tani, J. (2012). Spontaneous Prediction Error Generation in
1466   Schizophrenia.   PLoS One, 7(5): e37843.
1467
1468   Yamashita, Y., & Tani, J. (2008). Emergence of functional hierarchy in a multiple
1469   timescale neural network model: a humanoid robot experiment. PLoS Computational
1470   Biology, 4 (11), e1000220.
1471
1472