proceedings

Abstract

A New Look at Habits Using Simulation Theory†

Erik Billing[1]*

1 School of Informatics, University of Skövde; erik.billing@his.se
* Correspondence: erik.billing@his.se; Tel.: +46-500-448367

† Presented at the IS4SI 2017 Summit DIGITALISATION FOR A SUSTAINABLE SOCIETY, Gothenburg, Sweden, 12-16 June 2017.

Abstract: Habits as a form of behavior re-execution without explicit deliberation is discussed in terms of implicit anticipation, to be contrasted with explicit anticipation and mental simulation. Two hypotheses, addressing how habits and mental simulation may be implemented in the brain and to what degree they represent two modes brain function, are formulated. Arguments for and against the two hypotheses are discussed shortly, specifically addressing whether habits and mental simulation represent two distinct functions, or to what degree there may be intermediate forms of habit execution involving partial deliberation. A potential role of habits in memory consolidation is also hypnotized.

With the endeavor to explain how humans achieve an "inner world" underlying many cognitive abilities, Barsalou [1], Hesslow [2], and Grush [3] have proposed similar theories suggesting that the inner world stem from reactivation of cortical regions producing mental simulations of the agent's interactions with the environment. These so-called *simulation theories* take an embodied stance to cognitive science, criticizing symbolic and other dis-embodied forms of representations, while emphasizing human's ability to internalize the outer world.

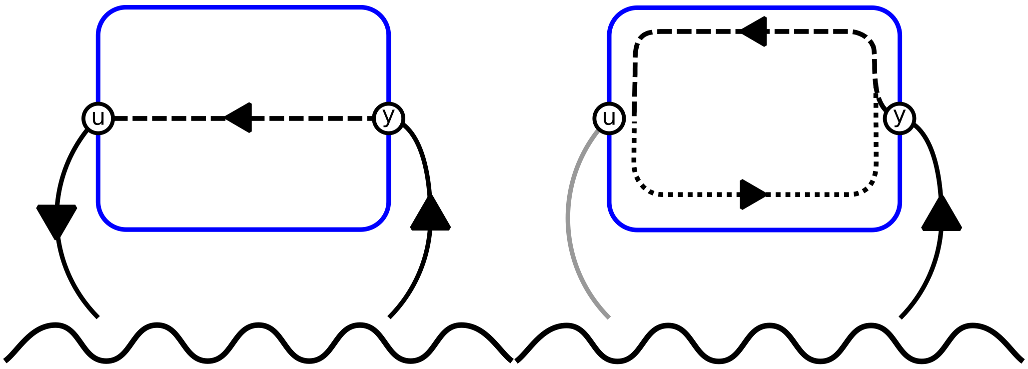

Figure 1. Left, an agent, encapsulated by a blue line, that perceives (y) and acts (u) upon its environment, depicted as a sine wave. The dashed line depicts the inverse model. **Right,** the same agent conducting internal simulation. Here, actions computed by the inverse model are inhibited (covert) and instead fed into a forward model (dotted line), computing expected percepts. As a result, the internal perception-action loop can continue without overt interactions with the world.

These internalized simulations are formulated as perception-action loops where the overt action is inhibited and replaced with a forward model anticipating the sensor-consequences of executing specific actions. The anticipated percepts can constitute input to an inverse model which produce new actions, iteratively generating long term predictions or simulations of the agent's interactions with the world. This process is depicted in Fig. 1.

While these theories are most frequently discussed in relation to mentalizing and planning, see e.g. Svensson [4] for a review, they are also relevant for other aspects of behavior formation. If the simulation theories are true, they entail re-activation or re-execution of behavior internally. The inverse model implementing the perception-action association is common for both overt (executed) and covert (simulated) behavior, and one key function of the internal simulation may be to update the inverse model in order to improve the agent's overt behavior.

Considering habits as an overt form of re-execution, we here explore the link between internal simulation and habits as two forms of covert and overt behavior re-execution, respectively. Indeed, Watson [5] discussed cognition and thinking in terms of motor habits as early as 1913 [6]. More recently, Pezzulo [7] discusses habits in terms of implicit anticipatory behavior, to be contrasted with explicit anticipation, i.e., mental simulation. With this view, habits rely on contextual activation of inverse models guiding the agent towards a goal (on-line), without any explicit representation of the goal state. This corresponds to the process depicted in Fig. 1, left. Mental simulation relies on similar contextual activation of inverse models, but requires in addition forward models computing the sensory consequences (off-line), eventually leading up to activation of rewarding stimuli that can be seen as an explicit representation of the goal state (Fig. 1, right).

Here, we draw on this link between habits and internal simulation in discussing an interesting question originally posed by Bubic et al. [8]: *Does implicit anticipation (habits) and explicit anticipation (simulation) entail a dichotomy or is it more appropriately discussed in terms of a continuous distribution of representations characterized by different degree of explicitness?* This question breaks down into at least two hypotheses:

H1: Habits and mental simulation depends on different systems in the brain.

H2: Habits and mental simulation constitute two distinct modes of mental function.

Relevant for this discussion is Downing's [9] review of predictive models in the brain. Downing associates implicit and explicit anticipation with different brain regions. Cerebellum and basal ganglia are, according to this review, primarily concerned with implicit anticipation, while neocortex, hippocampus, and thalamus has a larger role in explicit anticipation. Although habits are not directly discussed by Downing, this intuitively renders evidence supporting H1.

However, following the simulation theories and specifically the simulation hypothesis [2], mental simulation is argued to engage the same inverse models underlying implicit anticipation and habits, making a two systems view of anticipation more complex. An alternative reading of the separation suggested by Downing [9] is that cerebellum and basal ganglia are implementing inverse models, being the dominant regions during implicit anticipation but are likewise engaged in explicit anticipation. Neocortex, hippocampus, and thalamus would in this view do forward modeling, critical for explicit anticipation but less active during implicit anticipation and habit execution.

One consequence of the simulation theory view is that mental simulation can be well separated from habit execution, supporting H2, but may still play a key role in habit formation. Specifically, mental simulation produce simulated feedback that may constitute learning signals for both forward and inverse models, and hence affecting also habitual behavior. Evidence for this phenomenon can be drawn from studies of athletes and rehabilitation patients increasing performance as a result of motor imagery [10], [11]. However, studies investigating the specific relation between habits and imagery are to our knowledge absent.

Wood and Neal [12] discuss the relation between habits and goals. Following a classical two-process theory perspective referring to more cognitivistic, symbol based, formulation of a deliberative system [13], they make a clear distinction between habits and goal directed behavior. However, Wood and Neal also argues for two modes of queuing habits. One *direct* form, associating a specific context with responses, and one *motivated* form, associating the goal of a behavior with the present context. Motivated cuing is phrased in terms of "context cues […] carry hot motivational influence insofar as they signal opportunities to perform reward responses" [12]. Motivated cuing implies some form of goal representation and thereby, in our present discussion, at least partial engagement of explicit anticipation. As such, motivated cuing of habits could constitute an instance in-between implicit and explicit anticipation and with that, evidence against Hypothesis 2.

Computational models could be used as another way to investigate the relation between mental simulation and habits. In previous work [14], we analyze the behavior of the robot learning algorithm

Predictive Sequence Learning [15], controlling a robot during both overt and covert action. The overt mode corresponds to implicit anticipation and could be understood as habits in the context discussed here. Covert action implies internal simulation and make use of both inverse and forward generative models. By mixing contextual information for several different human demonstrations, the robot produces novel goal-directed behavior during both overt and covert modes. Although focus of the work was on covert action, the overt mode can be understood as a demonstration of novel goal-directed behavior without explicit deliberation. If this result holds also for mental simulation in humans, it highlights the importance of distinguishing goal representations in terms of explicit anticipation through simulation from contextual information associated with a goal, but lacking explicit goal-directed deliberation. In humans, this distinction becomes increasingly complex as simulations are likely to occur on many levels of abstraction [16], leaving the door wide open for partial deliberation in terms of simulations on higher levels of abstraction, without engaging lower levels closer to sensory-motor interactions.

While the literature comprises a range of computational models [17]–[19] inspired by the simulation theories, none has, to our knowledge, been directly analyzed in relation to habits. One interesting aspect to investigate would be the formation of habits in relation to mental imagery. Following the arguments made here, habits should be subject to change as a result of mental imagery deliberately introducing new behavior. Habits could possibly also be strengthened as a result of conscious or unconscious mental simulation.

Simulation theory could also provide an argument why such a large proportion of our time is spent executing habits [12]. It has been hypothesized [4], [20] that one reason for the brain engaging in unconscious mental simulation, e.g. in the form of dreaming, is memory consolidation. Many computational models of motor behavior, taking our own previous work as examples [14], [15], requires repeated presentation of training data to achieve stable generative models. If this phenomenon applies also to the brain, habits and dreaming could reflect forms of overt and covert memory consolidation, providing repeated presentations of similar sensory-motor relations necessary to stabilize internal models of the world.

References

1. L. W. Barsalou, "Simulation, situated conceptualization, and prediction," *Philos. Trans. R. Soc. B Biol. Sci.*, vol. 364, no. 1521, pp. 1281–1289, May 2009.

2. G. Hesslow, "Conscious thought as simulation of behaviour and perception," *Trends Cogn. Sci.*, vol. 6, no. 6, pp. 242–247, 2002.

3. R. Grush, "The emulation theory of representation: motor control, imagery, and perception.," *Behav. Brain Sci.*, vol. 27, no. 3, pp. 377–396, 2004.

4. H. Svensson, "Simulations," Linköping University, Linköping, 2013.

5. J. B. Watson, "Psychology as the behaviorist views it," *Psychol. Rev.*, vol. 20, pp. 158–177, 1913.

6. G. Hickok, "Eight Problems for the Mirror Neuron Theory of Action Understanding in Monkeys and Humans," *J. Cogn. Neurosci.*, vol. 21, no. 7, pp. 1229–1243, 2009.

7. G. Pezzulo, "Coordinating with the future: The anticipatory nature of representation," *Minds Mach.*, vol. 18, no. 2, pp. 179–225, 2008.

8. A. Bubic, D. Y. von Cramon, and R. I. Schubotz, "Prediction, cognition and the brain.," *Front. Hum. Neurosci.*, vol. 4, p. 25, Jan. 2010.

9. K. Downing, "Predictive models in the brain," *Conn. Sci.*, vol. 21, no. 1, pp. 39–74, 2009.

10. A. Guillot and C. Collet, "Duration of Mentally Simulated Movement: A Review," *J. Mot. Behav.*, vol. 37, pp. 10–20, 2005.

11. J. Munzert, B. Lorey, and K. Zentgraf, "Cognitive motor processes: The role of motor imagery in the study of motor representations," *Brain Res. Rev.*, vol. 60, no. 2, pp. 306–326, 2009.

12. W. Wood and D. T. Neal, "A new look at habits and the habit-goal interface," *Psychol. Rev.*, vol. 114, no. 4, pp. 843–863, 2007.

13. E. R. Smith and J. DeCoster, "Dual-Process Models in Social and Cognitive Psychology: Conceptual Integration and Links to Underlying Memory Systems," *Personal. Soc. Psychol. Rev.*, vol. 4, no. 2, pp. 108–131, 2000.

14. E. A. Billing, H. Svensson, R. Lowe, and T. Ziemke, "Finding Your Way from the Bed to the Kitchen: Reenacting and Recombining Sensorimotor Episodes Learned from Human Demonstration," *Front. Robot. AI*, vol. 3, no. March, p. 9, 2016.

15. E. A. Billing, T. Hellström, and L. E. Janlert, "Simultaneous Recognition and Reproduction of Demonstrated Behavior," *Biol. Inspired Cogn. Archit.*, vol. 12, pp. 43–53, 2015.

16. M. Jung, J. Hwang, and J. Tani, "Self-Organization of Spatio-Temporal Hierarchy via Learning of Dynamic Visual Image Patterns on Action Sequences," *PLoS One*, vol. 10, no. 7, pp. 1–16, 2015.

17. M. Haruno, D. M. Wolpert, and M. Kawato, "Mosaic model for sensorimotor learning and control.," *Neural Comput.*, vol. 13, no. 10, pp. 2201–2220, Oct. 2001.

18. T. Ziemke, D.-A. Jirenhed, and G. Hesslow, "Internal simulation of perception: a minimal neuro-robotic model," *Neurocomputing*, vol. 68, pp. 85–104, Oct. 2005.

19. Y. Demiris and B. Khadhouri, "Hierarchical attentive multiple models for execution and recognition of actions," *Rob. Auton. Syst.*, vol. 54, no. 5, pp. 361–369, May 2006.

20. S. Thill and H. Svensson, "The inception of simulation: a hypothesis for the role of dreams in young children," in *Proceedings of the 33rd Annual Conference of the Cognitive Science Society*, 2011, pp. 231–236.