

A model of Theory-of-Mind based on action prediction

Johannes Schultz^{1,2} and Frédérique de Vignemont³
 1. Wellcome Dept. of Imaging Neuroscience, University College London, UK
 2. Max-Planck Institute for Biological Cybernetics, Tübingen, Germany
 3. Institute of Cognitive Neuroscience, University College London, London, UK

1 - Introduction

Theory-of-Mind, or mentalising, is defined as a cognitive process used to understand other peoples' actions based on mental states. Two main theories of mentalising have been much discussed in recent years: Simulation Theory (e.g. Goldman 1993; Gallese and Goldman 1998), and Theory-Theory (e.g. Gopnik 1993). Models derived from these ideas and integrating neuroscience findings have been proposed since (e.g. Frith and Frith 1999; Blakemore and Decety 2001; Wolpert, Doya and Kawato 2003). Here we use the main idea from one of these models (Wolpert et al 2003) to propose a tentative model of mentalising based on action prediction and semantic representation. We also review a few neuroimaging studies of the processes involved in the model.

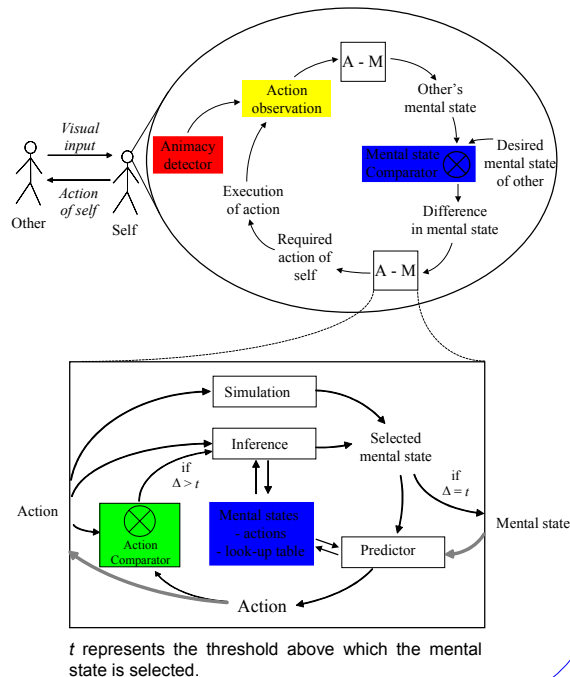
2 - The model

The model is proposed to work in the following way:

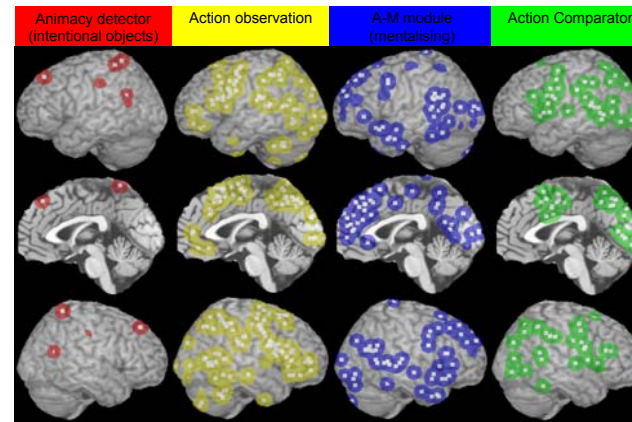
- 1) an **Animacy Detector** gets triggered in the presence of potentially animate agents,
- 2) an **Action Observation** module recognizes their actions,
- 3) the **A-M module** selects a potential mental states for these actions,
- 4) predict further actions of the agent (for mental state validation and adaptation of self), compared to real actions by the **Action Comparator**
- 5) **compare** the potential mental state to a desired mental state of the agent (e.g. during communication),
- 6) calculates actions to be taken by observer to induce change in mental state of the agent.

Inputs for the model are the visual aspect of the agent and (in case of an interaction with that agent) a representation of the desired mental state of the agent.

The A-M module works in both directions: associating a mental state to an observed action or predicting an action given a particular mental state (see grey arrows).



3 - Potential neural correlates



Here are reviewed neuroimaging studies that might correspond to 4 components of the model:

- 1) detection of animacy (studies used abstract objects moving so as to induce the percept of intentionality or animacy),
- 2) action observation (generally videos of hand, arm or full-body actions compared to a resting hand/arm/body),
- 3) mentalising (stories, cartoons, animations describing mental states compared to similarly complex items without mental states), and
- 4) Action comparison (e.g. an action or action sequence where the (last) action is different from the one predicted, or the equivalent with a sequence of abstract stimuli).

For references coded by color, see below

References

Theories and models of mentalising

- Blakemore SJ, Decety J. *Nat Rev Neurosci*. 2001 Aug;2(8):561-7.
 Frith CD, Frith U. *Interacting minds—a biological basis*. *Science*. 1999 Nov 26;286(5445):1692-5.
 Gallese, V. & Goldman, A. *Trends Cogn Sci*. 1998; 2:493-501.
 Goldman, A. *Consciousness and Cognition*. 1993; 2:364-382.
 Chaminade T, Decety J. *Neuroreport*. 2002 Oct 28;13(15):1975-8.
 Schultz J, Friston KJ, O'Doherty J, Wolpert DM, Frith CD. *Neuron*. 2005 Feb 17;45(4):625-35.
 Schultz J, Imamizu H, Kawato M, Frith CD. *J Cogn Neurosci*. 2004 Dec;16(10):1695-705.
 Schultz RT, Grelotti DJ, Klin A, Kleinman J, et al. *Philos Trans R Soc Lond B*. 2003 Feb 28;358(1430):415-27.

Moving "intentional" objects

- Blakemore SJ, Boyer P, Pachot-Clouard M, Meltzoff A, et al. *Cereb Cortex*. 2003 Aug;13(8):837-44.
 Castelli F, Happe F, Frith U, Frith C. *Neuroimage*. 2000 Sep;12(3):314-25.
 Chaminade T, Decety J. *Neuroreport*. 2002 Oct 28;13(15):1975-8.
 Schultz J, Friston KJ, O'Doherty J, Wolpert DM, Frith CD. *Neuron*. 2005 Feb 17;45(4):625-35.
 Schultz J, Imamizu H, Kawato M, Frith CD. *J Cogn Neurosci*. 2004 Dec;16(10):1695-705.
 Schultz RT, Grelotti DJ, Klin A, Kleinman J, et al. *Philos Trans R Soc Lond B*. 2003 Feb 28;358(1430):415-27.

Action observation

- Buccino G, Binkofski F, Fink GR, Fadiga L, et al. *Eur J Neurosci*. 2001 Jan;13(2):400-4.
 Decety J, Grezes J, Costes N, Perani D, et al. *Brain*. 1997 Oct;120 (Pt 10):1763-77.
 Grezes J, Armony JL, Rowe J, Passingham RE. *Neuroimage*. 2003 Apr;18(4):928-37.
 Grezes J, Costes N, Decety J. *Cognitive Neuropsychology*. 1998; 15 (6/7/8): 553-582
 Iacoboni M, Molnar-Szakacs I, Gallese V, Buccino G, et al. *PLoS Biol*. 2005 Mar;3(3):e79.
 Tettamanzi M, Buccino G, Saccurian MC, Gallese V, et al. *J Cogn Neurosci*. 2005 Feb;17(2):273-81.
 Saygin AP, Wilson SM, Hagler DJ Jr, Bates E, et al. *J Neurosci*. 2004 Jul 7;24(27):6181-8.

Mentalising

- Berthoz S, Armony JL, Blair RJ, Dolan RJ. *Brain*. 2002 Aug;125(Pt 8):1696-708.
 Brunet E, Sarfati Y, Hardy-Bayle MC, Decety J. *Neuroimage*. 2000 Feb;11(2):157-66.
 Castelli F, Happe F, Frith U, Frith C. *Neuroimage*. 2000 Sep;12(3):314-25.
 Ferstl EC, von Cramon DY. *Neuroimage*. 2002 Nov;17(3):1599-612.
 Fletcher PC, Happe F, Frith U, Baker SC, Dolan RJ, et al. *Cognition*. 1995 Nov;57(2):109-28.
 Gallagher HL, Happe F, Brunswick N, Fletcher PC, et al. *Neuropsychologia*. 2000;38(1):11-21.
 Goel V, Grafman J, Sadato N, Hallett M. *Neuroreport*. 1995 Sep 11;6(13):1741-6.
 Schultz RT, Grelotti DJ, Klin A, Kleinman J, et al. *Philos Trans R Soc Lond B*. 2003 Feb 28;358(1430):415-27.
 Schultz J, Friston KJ, O'Doherty J, Wolpert DM, Frith CD. *Neuron*. 2005 Feb 17;45(4):625-35.
 Schultz J, Imamizu H, Kawato M, Frith CD. *J Cogn Neurosci*. 2004 Dec;16(10):1695-705.
 Saxe R, Wexler A. *Neuropsychologia*. 2005;43(10):1391-9.
 Saxe R, Xiao DK, Kovacs G, Perrett DI, Kanwisher N. *Neuropsychologia*. 2004;42(11):1435-48.
 Vogtley K, Bussfeld P, Newen A, Herrmann S, et al. *Neuroimage*. 2001 Jul;14(1 Pt 1):170-81.

Action / intention incongruencies

- Downar J, Crawley AP, Mikulis DJ, Davis KD. *Nat Neurosci*. 2000 Mar;3(3):277-83.
 Downar J, Crawley AP, Mikulis DJ, Davis KD. *Neuroimage*. 2001 Dec;14(6):1256-67.
 Downar J, Crawley AP, Mikulis DJ, Davis KD. *J Neurophysiol*. 2002 Jan;87(1):615-20.
 Manthey S, Schubotz RI, von Cramon DY. *Brain Res Cogn Brain Res*. 2003 Feb;15(3):296-307.
 Pelphrey KA, Morris JP, McCarthy G. *J Cogn Neurosci*. 2004 Dec;16(10):1706-16.
 Pelphrey KA, Morris JP, McCarthy G. *Brain*. 2005 May;128(Pt 5):1038-48.
 Schubotz RI, von Cramon DY. *Cereb Cortex*. 2001 Mar;11(3):210-22.

Acknowledgements

Thanks to Chris D. Frith for numerous discussions, comments and encouragements.