Agents, Actions, and Social Reality

Yasuo NAKAYAMA
Graduate School of Human Sciences, Osaka University
nakayama@hus.osaka-u.ac.jp

In this presentation, we propose *BDOI*-model of atomic agents. *BDOI*-model characterizes mental states of an atomic agent through triple ⟨*belief*, *desire*, *normative belief*⟩ and explains actions based on mental states and intention. Parts of mental states and interpretation of terms can be shared among atomic agents. An aim of this presentation is to explain the construction of social reality based on an analysis of agents, shared mental states, and actions.

1. Model of Atomic Agents

Donald Davidson provided the standard theory of action (Davidson 1980). The core of this theory can be expressed by the following three theses (Schlosser 2021, Sect. 2 and Sect. 2.1).

(1a) The notion of *intentional action* is more fundamental than the notion of *action*.
(1b) There is a close connection between intentional action and acting for a reason.
(1c) [Theory of agency] A being has the capacity to act intentionally just in case it has the right functional organization: just in case the instantiation of certain mental states and events (such as desires, beliefs, and intentions) would cause the right events (such as certain movements) in the right way.

In this presentation, we accept the first two theses and modify (1c). We propose that mental states can be characterized by beliefs, desires, and normative beliefs and that an intention leads an agent, based on the given mental states, to performance of an action. John Searle pointed out that an action may have *desire-based reason* or *desire-independent reason* (Searle 2010: Chap. 6 Sect. 1). Modifying Searle′s position, we propose that an action may have reason that is based on both desires and normative beliefs.

Now, we start our proposal with the following description of atomic agents and call it *BDOI-model of atomic agents* (Nakayama 2017a, 2021).

(2a) [Atomic agent] An atomic agent can perform some actions. A part of mental states of an atomic agent can be characterized through triple ⟨*belief*, *desire*, *normative*

belief⟩, where each of the three components of the triple is a set of First-Order sentences (*FO*-sentences). I call this triple *BDO*-system. This *BDO*-system can be updated when an agent obtains new information.

(2b) [Transparency] Mental states described in (2a) is transparent in the following sense:

(i) [Belief] If *A* believes that φ, then *A* knows that *A* believes that φ.

(ii) [Desire] If *A* desires that φ, then *A* knows that *A* desires that φ.

(iii) [Normative belief] If *A* believes that it is obligated that φ, then *A* knows that *A* believes that it is obligated that φ.

(2c) [Intention as decision making] An *atomic agent* chooses an action type based on her/his *BDO*-system and performs it. In such a case, we say that this agent *intentionally performed* this action.

We use **not**, **&**, ***or***, ⇒, and ⇔ as meta-language expressions of logical connectives. Pair ⟨*BB*, *OB*⟩ which is a subsystem of *BDO*-system ⟨*BB*, *DB*, *OB*⟩ is called a *BO-system*. Let *cons* be an abbreviation of consistent and $Cn(X)$ be an abbreviation of *the deductive closure of X*.

(3a) [Belief] $\mathbf{B}_{BDO}\,\varphi \Leftrightarrow_{\text{def}} (cons(BB)\ \&\ \varphi \in Cn(BB))$

(3b) [Possibility] $\mathbf{M}_{BDO}\,\varphi \Leftrightarrow_{\text{def}} cons(BB \cup \{\varphi\})$

(3c) [Obligation] $\mathbf{O}_{BDO}\,\varphi \Leftrightarrow_{\text{def}} (cons(BB \cup OB)\ \mathbf{\&}\ \varphi \in Cn(BB \cup OB)\ \mathbf{\&}\ \textbf{\textit{not}}\ (\varphi \in Cn(BB)))$

(3d) [Prohibition] $\mathbf{F}_{BDO}\,\varphi \Leftrightarrow_{\text{def}} \mathbf{O}_{BDO}\,\neg\varphi$

(3e) [Permission] $\mathbf{P}_{BDO}\,\varphi \Leftrightarrow_{\text{def}} (cons(BB \cup OB \cup \{\varphi\})\ \mathbf{\&}\ \textbf{\textit{not}}\ (\varphi \in Cn(BB)))$

(3f) [Desire] $\mathbf{D}_{BDO}\,\varphi \Leftrightarrow_{\text{def}} (cons(BB \cup DB)\ \mathbf{\&}\ \varphi \in Cn(BB \cup DB)\ \mathbf{\&}\ \textbf{\textit{not}}\ (\varphi \in Cn(BB)))$

(3g) *BDO* is consistent $\Leftrightarrow_{\text{def}} (cons(BB \cup OB)\ \mathbf{\&}\ cons(BB \cup DB))$

(3h) [Respect] Atomic agent *A* with *BDO*-system ⟨*BB*, *DB*, *OB*⟩ respects *BO*-system ⟨BB_s, OB_s⟩ $\Leftrightarrow_{\text{def}}$ $BB_s \subseteq BB$ & $OB_s \subseteq OB$ & any action type that *A* chooses to perform is compatible with $BB_s \cup OB_s$.

According to (3h), an atomic agent who respects a *BO*-system obeys any obligation in the *BO*-system and she/he chooses only action types that are permitted in the *BO*-system. For example, a player of chess respects the *BO*-system of chess and she/he plays chess keeping out of violation of the *BO*-system.

We can update a *BDO*-system ⟨*BB*, *DB*, *OB*⟩ by updating *BB* or *DB* or *OB*. We call the framework that allows this kind of updates *Dynamic BDO-Logic*. A *BDO*-system in *Dynamic BDO-Logic* contains information about its stage. We write a *BDO*-system of *Dynamic BDO-Logic* as follows: $BDO(k) = \langle BB(k), DB(k), OB(k)\rangle$. A play of standard

two-man games can be described in *Dynamic BO-Logic* that is a subsystem of *Dynamic BDO-Logic* (Nakayama 2016, 2017a, 2021).

According to Searle, there are two types of rules, namely *regulative* and *constitutive rules* (Searle 1969: Chap. 2.5). Regulative rules regulate a pre-existing activity, an activity whose existence is logically independent of the rules. Regulative rules characteristically take the form of or can be paraphrased as imperatives, e.g., "Officers must wear ties at dinner". Constitutive rules constitute an activity the existence of which is logically dependent on the rules. Constitutive rules can be paraphrased as "*X* counts as *Y* in context *C*". A typical example is an introduction of a term used in a game., e.g., "A checkmate is made when the king is attacked in such a way that no move will leave it unattacked" (p. 34f). Both rules can be expressed in *BO-Logic*. In *BO*-system for officers $\langle BB_{officer}, OB_{officer}\rangle$, the *FO*-translation of sentence "Officers wear ties at dinner" is a member of $OB_{officer}$. Similarly, in *BO*-system of chess $\langle BB_{chess}, OB_{chess}\rangle$, the *FO*-translation of sentence "A checkmate is made if and only if the king is attacked in such a way that no move will leave it unattacked" is a member of BB_{chess}.

2. Ontology for Actions and Agents

Davidson developed an event ontology and considered events as First-Order objects as well as things. Furthermore, he interpreted actions as events that are intentional under some descriptions (Davidson 1980). Nakayama (2017b, 2019) extended this event-based semantics of Davidson and developed an axiomatic theory for *Four-Dimensional Event Ontology* (4EO). This theory is based on *General Extensional Mereology* (GEM) for (four-dimensionally extended) events. 4EO claims that everything is a four-dimensional object (4D-object).

(4a) The *universe* is the maximal 4D-object. This means that any 4D-object is a part of the universe.
(4b) An *event* is a 4D-bject. Thus, an action is also a 4D-object.
(4c) An *agent* is a 4D-object.

Atomic agents can share some parts of their mental states. We describe shared mental states of a group of atomic agents as follows.

(5a) Let group G be the mereological sum of atomic agents $A_1, \ldots, A_n$. Let $BDO(A_k) = \langle BB(A_k), DB(A_k), OB(A_k)\rangle$.
(5b) [Shared belief] φ is a shared belief in $G \Leftrightarrow_{def}$ for all A_k in G, $\varphi \in BB(A_k)$.

(5c) [Shared desire belief] φ is a shared desire belief in G $\Leftrightarrow_{def}$ for all A_k in G, $\varphi \in DB(A_k)$.

(5d) [Shared obligation belief] φ is a shared obligation belief in G $\Leftrightarrow_{def}$ for all A_k in G, $\varphi \in OB(A_k)$.

(5e) [Shared interpretation] All agents in G share interpretation of language L $\Leftrightarrow_{def}$ every agent in G interprets all symbols in L in the same way.

(5f) [Shared *BO*-system] *BO*-system $\langle BB, OB \rangle$ is shared in G $\Leftrightarrow_{def}$ all atomic agents in G share all beliefs in *BB*, all obligation beliefs in *OB*, and interpretation of all symbols in *BB* and *OB*.

(5g) [Game players] If G is a group of players of a game that is defined by a *BO*-system, then this *BO*-system is shared in G and respected by all players in G.

Now, the notion of *extended agent* can be specified as follows (Nakayama 2013).

(6a) [Atomic agent] An *atomic agent* is an agent. Any spatial part of an atomic agent is no agent.

(6b) [Agents and tools] Let *temporal-part* (x, t) denote the temporal part of object x in extended time t. Let A be an agent who uses thing B in t to perform an action. Then, the mereological sum *temporal-part* (A, t) + *temporal-part* (B, t) is an agent.

(6c) [Collective action] For every agent A who is a part of G, if E is a collective action performed by G, then there is an action of A that is a part of E.

(6d) If group G of agents performs a collective action in t, then *temporal-part* (G, t) is an agent.

(6e) If an object satisfies neither (6a) nor (6b) nor (6d), then it is no agent. (Note that this definition of *action* is recursive.)

(6f) [Extended agent] An agent that is not atomic is called an *extended agent*.

The collectivity is created based on the ability of people to share parts of mental states and interpretation of a language. In general, an extended agent is more than the fusion of atomic agents, because it can contain several artifacts as its components (see (6b)). If B_1 is the building of a factory, M_1 is the machine in B_1, and A_1, … , A_n are workers in B_1, and t denotes working hours, then *temporal-part* $((A_1 + \ldots + A_n) + M_1)$, t) is an extended agent. The workers in B_1 produce goods with M_1 and this production is a collective action (see (6c)). It is a characteristic of our description of collective actions that it takes artifacts as well as humans into consideration.

3. Social Actions and Social Reality

Max Weber thought that social actions of individuals construct the society. Thus, Weber characterized sociology as a science which attempts the interpretive understanding of social action to arrive at a casual explanation of its course and effects (Weber 1922: Sect. 1). This proposal looks persuasive, but it is also true that the society supports social actions. This means that the society and social actions are interconnected. Searle pointed out that some action types and some mental states presuppose some *social institutions*. For example, you can desire to have much money and buy things with money, because there is a monetary system established in the society (Searle 2010: Chap. 6, Sect. 1). This monetary system can be interpreted as *BO*-system $\langle BB_{ms}, OB_{ms}\rangle$ that is shared and respected by almost all members of this society.

Many actions presuppose the existence of the society. For example, if you use a smart phone to play a game, you need a smart phone that is invented and produced in the past. Based on this invention and the spread of smart phones, the action type of *using a smart phone* is created. This type of creation presupposes shared beliefs and shared interpretations of terms for some artifacts. Another type of creation can be found in games. For example, *hitting a home run* is particular action type in a baseball game. This type of creation presupposes shared *BO*-system for a game and shared interpretation of terms in the *BO*-system. Additionally, *playing a team game* presupposes some shared desires among members of a team. These examples show that the existence of many current actions presupposes some current shared *BO*-systems and some collective actions in the past.

References

Davidson, D. (1980) *Essays on Actions and Events*, Oxford University Press.

Nakayama, Y. (2013) The Extended Mind and the Extended Agent, *Procedia Social and Behavioral Sciences*, vol. 97, Elsevier, pp. 503-510.

Nakayama, Y. (2016) Norms and Games as Integrating Components of Social Organizations, H. Ishiguro *et al*. (eds), *Cognitive Neuroscience Robotics* B, Springer, pp. 253-271.

Nakayama, Y. (2017a) Philosophical Basis for Dynamic Belief-Desire-Obligation Logic, *Kyoto Philosophical Logic Workshop III 2017. 9. 7-9, Kyoto University*. https://drive.google.com/file/d/0Bz6DSBj3ixUZU2Q0Q0gzaDN4Mnc/view

Nakayama, Y. (2017b) Event Ontology based on Four-dimensionalism, *Department Bulletin Paper,* Graduate School of Human Sciences, Osaka University, vol. 43, pp.

175-192, http://hdl.handle.net/11094/60581

Nakayama, Y. (2019) *From Philosophy of Language to Metaphysics: A new Development in Four-dimensionalism*, Keiso shobo, in Japanese.

Nakayama, Y. (2021) *Living Together in Society: Philosophy of Society and Ethics*, Keiso shobo, in Japanese.

Schlosser, M. (2021) Agency, E.N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy* (Winter 2021 Edition).

Searle, J.R. (1969) *Speech Acts: An Essay in the Philosophy of Language*, Cambridge University Press.

Searle, J.R. (2010) *Making the Social World: The Structure of Human Civilization*, Oxford University Press.

Weber, M. (1922) *Soziologische Grundbegriffe*, in: *Wirtschaft und Gesellscaft* (1922), J. C. B. Mohr.