Logics of agency

Chapter 1: Introduction to action and agency

Daniele Porello and Nicolas Troquard

KRDB Research Center for Knowledge and Data
Free University of Bozen-Bolzano

name.surname@unibz.it

EASSS 2018

Overview of this chapter

Towards a logic of agency:

- Action and Agency.
- What is an action? (granularity, intentionality, deliberation)
- Actions and events.
- The intentional stance.
- Individual and collective actions.
- Natural language analysis of agentive sentences.
- A very short history of the logical analysis of agency.
- The logic of agency.
- Overview of this course.

Some notions presented in this course

- action, acting, doing an action,
- choice, making a choice, making a deliberative choice,
- being responsible for a result,
- obligations to do,
- having the capacity of making a choice,
- ...
- using tools,
- consuming resources, producing resources

Agent and Agency

Agent: an entity typically with some mental state, that is capable of action

Agency: exercise or manifestation of a capacity of action

Basic actions

> *A basic action is perfectly simple in the same sense in which the old 'simple ideas' were said to be: they were not compounded out of anything more elementary than themselves, but were instead the ultimately simple elements out of which other ideas were compounded. [Danto 65]*

Some challenges of basic action in knowledge representation.

- "Raise your hand":
  - "Are you the chairperson of this assembly or a choreographer?"
- Avrel! Kill this coyote!
  - Should I shoot it or strangle it, Joe?

Problems/difficulties:

- granularity of action
- meaningful distinction between actions

Intentionality

- Action is explained by intentionality [Anscombe 1957] [Davidson 2001]
  - one event can be more than one action under different descriptions
  - some actions are generated by other actions (see also [Goldman 1970], [Danto 1986])
- Intentionality and acting for a reason come hand in hand

Action as a kind of event

Often, an action is considered a particular kind of event.

Action sentences are seen as denoting some logical combinations of relations between particulars. [Davidson 1967, 2001 Essay 6]

"John buttered the toast slowly with a knife" becomes:

$$\exists e(\textbf{butter}(e, \textit{John}, \textit{the toast}) \,\&\, \textbf{slowly}(e) \,\&\, \textbf{with a knife}(e))$$

where e is a variable, *John* and *the toast* are constants, and **butter**/3, **slowly**/1, and **with a knife**/1 are predicates (modifying clauses).

By means of this formalisation, one can validate the reasoning step:

if John buttered the toast slowly with a knife,
then John buttered the toast.

Avrel kills the coyote in many ways: how many actions?

$$\exists e(\textbf{kill}(e, \textit{Avrel}, \textit{the coyotte}))$$

$$\exists e(\textbf{kill}(e, \textit{Avrel}, \textit{the coyotte}) \,\&\, \textbf{strangling}(e))$$

$$\exists e(\textbf{kill}(e, \textit{Avrel}, \textit{the coyotte}) \,\&\, \textbf{shooting at}(e, \textit{the coyotte}))$$

$$\exists e(\textbf{kill}(e, \textit{Avrel}, \textit{the coyotte}) \,\&\, \textbf{shooting at}(e, \textit{the shrub}))$$

Immediate criticism

[...] the dominant logical template takes an agent as a wart on the skin of an action, and takes an action as a kind of event. [Belnap et al. 2001]

Doing deliberately

"John buttered the toast slowly with a knife deliberately."

Davidson argues that one cannot treat "deliberately" like other modifying clauses.

He proposes: the construction "It was intentional of *x* that *p*" where *x* names the agent and *p* is a sentence that says the agent did something. [Davidson 67]

E.g.: "Oedipus intentionally sought the slayer of Laius."
becomes
"It was intentional of Oedipus that he (himself) sought the slayer of Laius."

The intentional stance

> *"Hitting a ball is an action, falling down a flight of stairs is not. A theory of action seeks, among other things, to explain the distinctions we make." [Segal 1991]*

What makes an entity an acting entity of an event is generally acknowledged to be the intentionality in action (or the intention for an achievement). To be the agent of an event, one has to make a deliberate decision governed by one's beliefs and desires (e.g., [Anscombe 63], [Dennett 71]).

Bratman ([Bratman 86]) built upon this and proposes that intentions operate like a filter over every action in order to select the actions that are desired and believed to be successful.

In logic, e.g.: [Cohen & Levesque 90] [Rao & Georgeff 91].

Often, agency $\sim$ free will: something is caused by a deliberate action.

> *"The main weakness in treating agency as a synonym for free will is that such an approach ignores or only gives lip service to the social nature of agency and the pervasive influence of culture on human intentions, beliefs, and actions." [L.M. Ahearn 2001]*

The agent of the action

Who is the agent who is responsible for acting?

Individual vs Collective action: generally, the result of some deliberate collective decision via deliberation, argumentation, voting, ...

- Formation of shared goals and shared intentions (e.g., [Tuomella & Miller 1988], [Tuomela 95]).
- agents' "we-intention" as part of a group are built from individual attitudes
- many similarities with shared and common knowledge ([Tuomela & Balzer 1997])
- (rebuttal by [Searle 1990])
- Group agents, Group Agency, [List & Pettit 2011], [PBF14].

In logic, e.g.: [Dunin-Kęplicz & Verbrugge 2010], [Por18].

The social structure of action: complex/institutional action

Goldman's event generation [Goldman 70].

- distinction between physical events and institutional events
- agents act upon the physical environment
- generate some more physical events by a causal relation,
- generate institutional events by a conventional relation (similar to Searle's notion of 'counts-as' [Searle 1995]).

In logic, e.g.: [Grossi et al. 2008], [Herzig et al. 2011].

Intentions are sometimes too demanding

- Some entities are capable of action but do not possess representational mental states. ([Davidson 1982]: only human agents have the relevant mental attitudes because they require linguistic competence.)
  - “The team is putting a tremendous effort.”
  - “The dog bites the postman.”
  - “The computer is playing Go.”
- Some instances of human agency tend to be explained without the ascription of representational mental states.
  - “Alan hit Bernard in the nose as he suddenly turned around.”
  - “As he parked his car, Charlie woke up from his stupor. He had been driving without thinking about it.”

Agentivity of verb phrases

In [Belnap and Perloff 1988] (see also [Kenny 1963, Chap. VII]):

The sentence φ marks the agentiveness of agent *a* just in case φ may be usefully paraphrased as "*a* sees to it that φ".

Example: "Ann eats sand" marks the agentivity of Ann in some context iff we can substitute it with

"Ann sees to it that she eats sand"

in the context.

We will study the logical forms of theses sentences in details.

Natural language analysis and agentive modalities

Examples:

Ahab found the White Whale (1)

The Pequod sailed in search of Moby Dick (2)

Ahab sailed in search of Moby Dick (3)

Ishmael sailed in search of Moby Dick (4)

Belnap, N., Perloff, M., and Xu, M., 2001, *Facing the future*, Oxford: Oxford University Press.

BELNAP'S *agentives*

Belnap discusses linguistic forms to express actions, that appears to behave as operators on propositions:

I . brings it about that

II . makes it the case that

III . causes it to be the case that

IV . is responsible for the fact that

V . lets it be the case that

VI . allows it to be the case that

VII . takes steps in order that

VIII . behaves so that in consequence

IX . sees to it that

And selects "sees to it that" as the less ontologically or epistemologically committed with other concepts.

LINGUISTIC FORM OF AGENTIVE SENTENCES

Belnap proposes a restricted fragment of natural language to express the form of agentive sentences:

$$\alpha \text{ sees to it that } \varphi \tag{5}$$

where α is the name of an agent and φ a proposition.

$$[\alpha \textit{ stit} : \varphi] \tag{6}$$

- Es. *Ahab sailed in search of Moby Dick* can be translated to *Ahab saw to it that Ahab sailed in search of MobyDick*.
- Es. *Ahab found the White Whale* can be translated to *Ahab saw to it that Ahab found the White Whale*

Belnap et al, *Facing the Future*, Capitolo 1.

Agentivity test

$$\varphi \text{ is agentive for } \alpha \text{ iff } \varphi \text{ is equivalent to } [\alpha \textit{ stit} : \varphi] \tag{7}$$

Examples:

1. *Ahab sailed in search of Moby Dick* is agentive for Ahab?
2. *Ahab sailed in search of Moby Dick* is agentive for Ismael?
3. *Ahab found the White Whale* is agentive for Ahab?

Belnap et al, *Facing the Future*, Chapter 1.

A BIT OF HISTORY OF THE LOGICAL ANALYSIS OF AGENCY

- Anselm of Canterbury
- von Wright
- Chellas
- Pörn ed Elgesem.
- Belnap

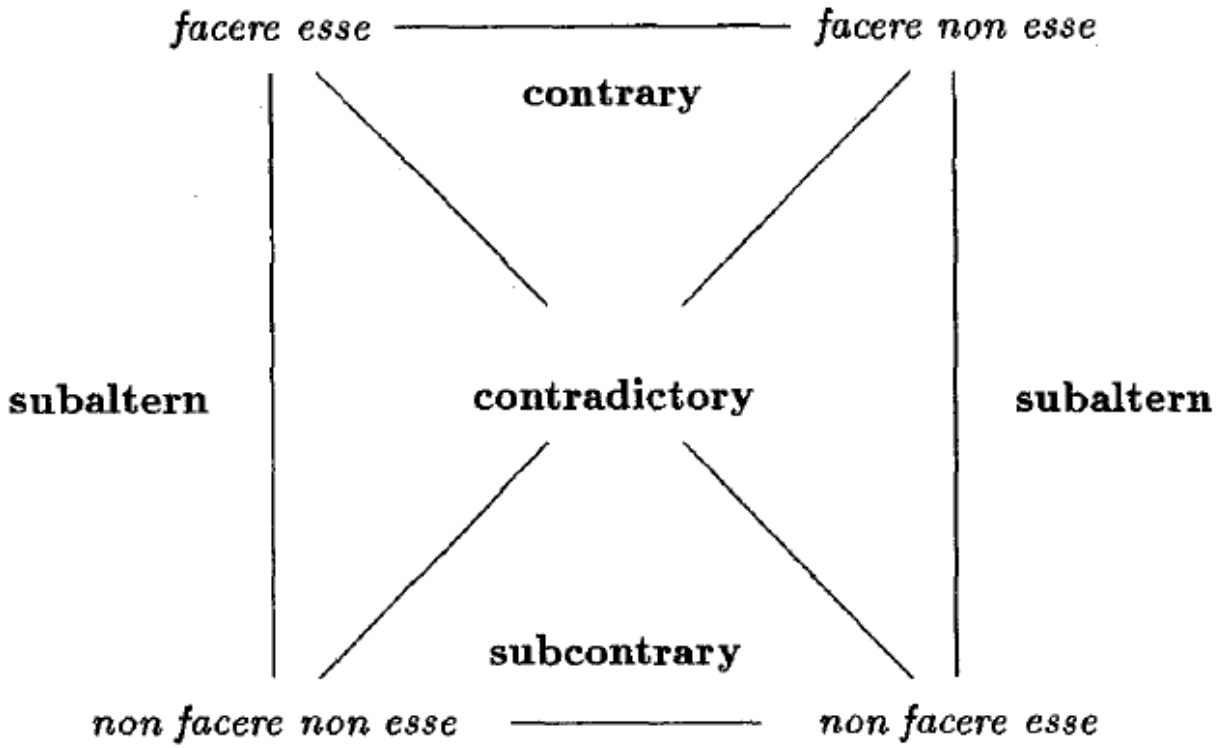
K. Segerberg. “Getting started: Beginnings in the logics of action”, Studia Logica 51, 347–378, 1992

Anselm of Canterbury (1033-1109)

- Anselm studies natural language forms such as *facere esse*, *facere non esse*, and he views them as operators for propositions:
- E.g. Henry interprets *facere esse* of Anselm as follows:
  - *facere esse* – *i* sees to it that *p*,
  - *facere non esse* – *i* sees to it that not *p*,
  - *non facere esse* – *i* does not see to it that *p*,
  - *non facere non esse* – *i* does not see to it that not *p*.

Paul Desmond Henry. *The logic of St. Anselm*, Claredon Press, Oxford, 1967.

A square of opposition for agentive sentences:



- von Wright is interested in the logic of actions as a part of deontic logic (e.g.“it is obligatory that φ” means “it is obligatory *doing* φ)
- “Action logics as a basis for deontic logic” (in *Practical Reasons*)
- Contra Davidson, von Wright argues that actions are not types of events.

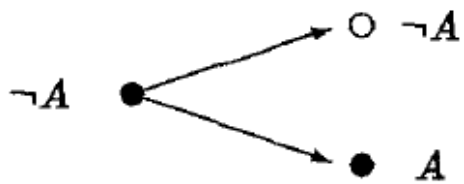
> *It would not be right to call acts a kind or species of events. An act is not a change in the world. But many acts may quite appropriately be described as the bringing about or effecting (’at will’) of a change. [von Wright 1963.]*

G. H. von Wright. *Norm and Action. A logical Inquire*. Routlege, 1963.
G. H. Von Wright, *Practical reason*, Blackwell, Oxford 1983

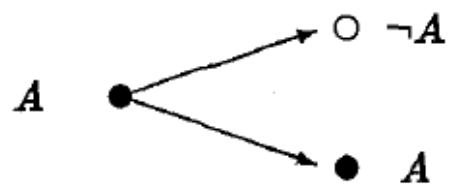
- von Wright lists 4 modes of relation between an agent and a state of affairs (represented by a proposition) φ.
- An agent can *produce*, *maintain*, *destroy*, *suppress* the state that φ.

- $\mathcal{B}\varphi$, the agent produces the state φ
- $\mathcal{S}\varphi$, the agent maintains the state φ
- $\neg\mathcal{B}\varphi$, the agent destroys the state φ
- $\mathcal{S}\neg\varphi$, the agent suppresses that state φ.

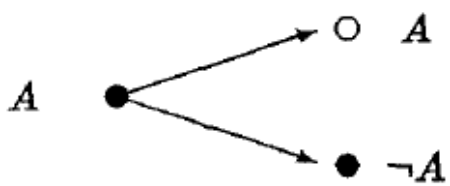
An hint to von Wright intended semantics:



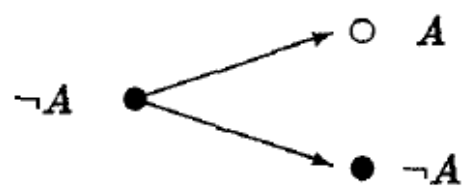
Producing A



Sustaining A



Destroying A



Suppressing A

Chellas developed the first modal logic of agency with an explicit formal modern semantics:

- Chellas formalises "*i* sees to it that φ" by $\Delta_i\varphi$.
- The semantics consists in a set of temporal states T and a set of possible states of the world S.
- A history is a function $h : T \to S$.

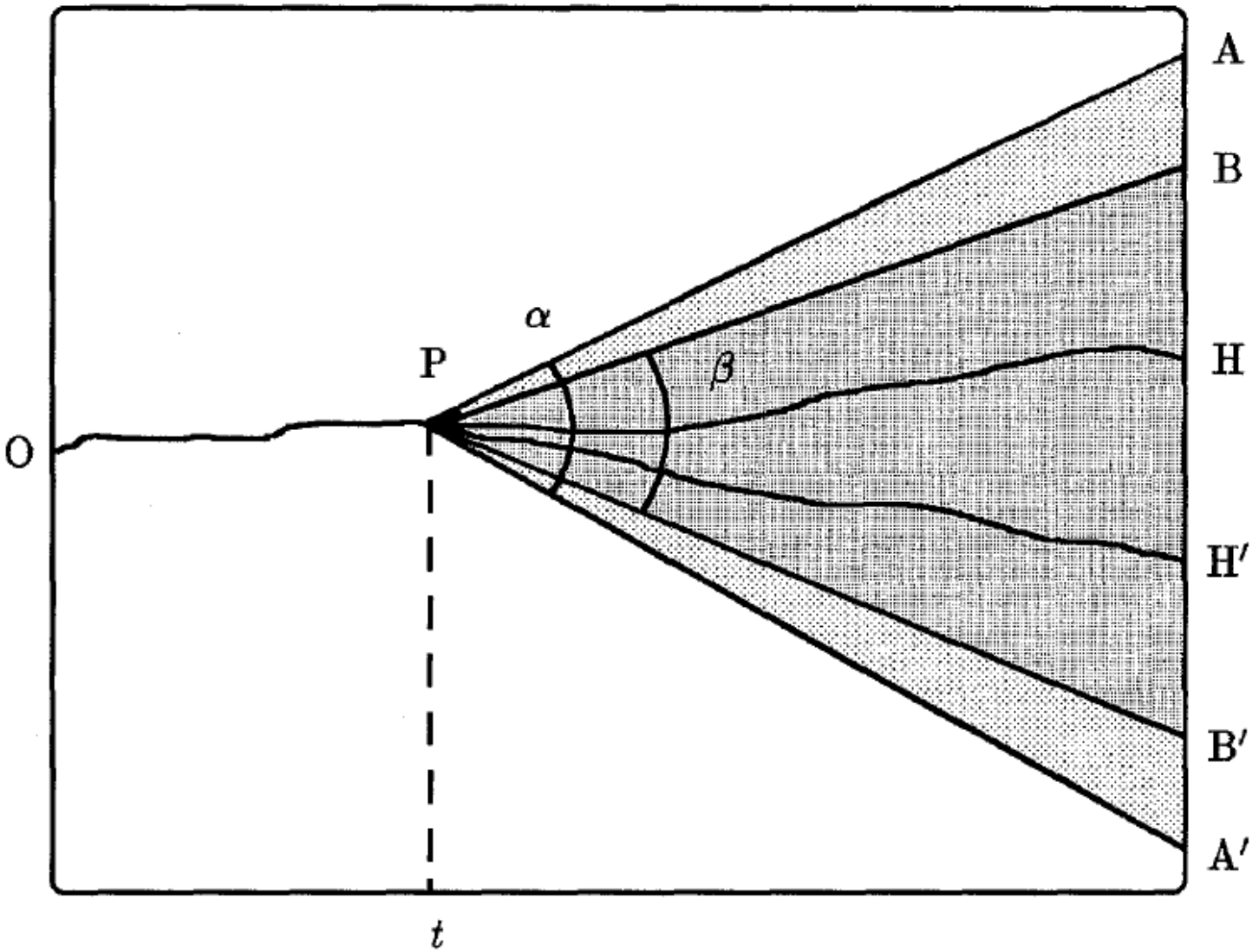
$$(h, t) \models \varphi \tag{8}$$

" φ is true respect to the history h ad moment t".

Chellas, B. (1969). The Logical Form of Imperatives. Perry Lane Press. (PhD Thesis, under the supervision of D. Scott)

CHELLAS STIT

A sketch of STIT semantics:



Belnap: STIT

- The STIT logic by Belnap extends the analysis of Chellas.
- The concept of action in STIT is based on two assumptions:
  1. An agent cannot determine a unique possible future by acting.
  2. An agent can prevent a future to be the case by acting.
- Ex. If I do eat an ice cream, I can exclude all the future where I do not eat an ice cream, but I cannot determine a unique future, e.g. the one where I eat an ice cream and all the other people do not.

Belnap, N., Perloff, M., and Xu, M., 2001, *Facing the future*, Oxford: Oxford University Press.

Pörn, Elgesem: The principle of agentivity

- $E_i\varphi$: "agent *i* brings about φ".
- Pörn ed Elgesem propose four principle of the modality "bringing it about":

1. If something is brought about, then this something holds.
2. It is not possible to bring about a tautology.
3. If an agent brings about two things concomitantly then the agent also brings about the conjunction of these two things.
4. If two statements are equivalent, then bringing about one is equivalent to bringing about the other.

Pörn, I. (1977). Action Theory and Social Science: Some Formal Models. Synthese Library 120. D. Reidel, Dordrecht.
Elgesem, D. (1997). The modal logic of agency. Nordic J. Philos. Logic, 2(2).
Governatori, G. and Rotolo, A. (2005). On the Axiomatisation of Elgesem's Logic of Agency and Ability. *Journal of Philosophical Logic*, 34:403Ð431.

- 1. $E_i\varphi \rightarrow \varphi$
- 2. $\neg E_i \top$ (dove $\top$ è una tautologia)
- 3. $E_i A \wedge E_i B \rightarrow E_i(A \wedge B)$
- 4. Se $\vdash A \leftrightarrow B$, allora $\vdash E_i A \leftrightarrow E_i B$.

To obtain a modal logic that satisfies 1–4, we shall abandon *normal* modal logics.

Why a modal logic of agency?

- A modal logic of agency interpret agentive statements as modal formulas where a modality $STIT_i$, E_i, DO_i, ..., operates on a proposition φ.
- The assumptions about the principle of agentivity can be studied by formalising the principles of those modalities.
- The aim is to provide the general logical principles that do not depend on a specific ontological view of what is an action.

Why a modal logic of agency?

- You can check whether your principles of agency are coherent (consistent).
- You can check whether your principles of agency are redundant.
- You can check the models of your principles of agency and possibly find counter examples.
- You can define a notion of reasoning about actions (i.e. a logic of practical reasoning).
- You can check whether your logic of agency is decidable.
- You can investigate the class of theorems (logical consequence) of your principles of agency.
- You can investigate the computational complexity of your logic.
- You can study the integration of your logic with deontic, epistemic, doxastic, ... logics.

A CRITIQUE TO THE MODAL LOGIC OF AGENCY

Kenny 1975, in normal modal logic, we have that:

$$\Diamond(A \vee B) \vdash \Diamond A \vee \Diamond B \tag{9}$$

- Kenny argues that $\Diamond$ is not apt to capture the *ability can*, which is essentially involved in *doing*.
- E.g. an agent is able to pick a red or a black card from a deck. But she or he is not able to pick a red card, and she or he is not able to pick a black card. Violating (9).
- Kenny concludes that $\Diamond$ is apt to express the *opportunity* "can" non the *ability* can. Therfore, $\Diamond$ is not apt to model actions.

if "John buttered the toast slowly with a knife", *then* "John buttered the toast"

This argument works for *normal* modal logic, not for weaker systems.

A. Kenny, *Will, freedom and power*, Blackwell, Oxford 1975.

Two traditions in the logic of action

In the literature on the logics of action, two types of modelling traditions:

- Logic of *agency* (Chellas, Belnap, Pörn, Elgesem, von Wright, and others).
- *Dynamic logics*.

Dynamic Logic

- Dynamic logics have been developed to model actions of programs.
- "after the execution of program a, φ holds".
- Pratt introduces two operators:
  - $[a]\varphi$: after every computation of a, φ holds.
  - $< a > \varphi$: after some computation of a, φ holds.
- Composition of programs e.g. ";" (sequential), $[a; b]\varphi$

V. R. Pratt, Application of modal logic to programming, Studia logica 39 (1980), pp. 257 - 274.

Belnap on Dynamic Logics

Belnap excludes Dynamic Logics from the true logic of agency because:

- Dynamic logics are about actions of programs, not of agents, so they are actually modelling "processes" rather than actions.
- There are names for actions in the language (es. a, b, $a; b$), thus there is an ontological commitment about actions.

Segerberg (1992) however, discusses Dynamic Logics within the tradition of the logics of action.

Applications in AI and MAS

- Modelling agency and practical reasoning are important in AI, KR, and MAS, as agency is a crucial aspect of the intelligent behaviour (in simulation, planning, and representation).
- E.g. BDI (belif-desire-intention) model of an intelligent agent (es. robots, software, sensors, ... whatever we want to construe as an agent) requires understanding and modelling actions (e.g. goals).

Stanford Enciclopedia of Philosophy, "Logic of Action"
M. J. Wooldridge, An Introduction to MultiAgent Systems, Chichester: Wiley. 2002.

Overview of this course

- Basics of dynamic logics.
- Belnap STIT.
- Non-normal (minimal) modal logics.
- The logic of *Bringing it About* (BIAT) (Pörn, Elgesem, Governatori and Rotolo, 2005).
- Recent topics: resource sensitive agency, based on [BPT14], [PT14], [PT15].

References I

Laura M. Ahearn.

Language and agency.
Annu. Rev. Anthropol., 30:109–137, 2001.

G. Elizabeth M. Anscombe.

Intention.
Basil Blackwell, 1957.

Robert Audi.

Acting for reasons.
Philosophical Review, 95(4):511–546, 1986.

Nuel Belnap and Michael Perloff.

Seeing to it that: a canonical form for agentives.
Theoria, 54(3):175–199, 1988.

Stefano Borgo, Daniele Porello, and Nicolas Troquard.

Logical operators for ontological modeling.
In *Formal Ontology in Information Systems - Proceedings of the Eighth International Conference, FOIS 2014, September, 22-25, 2014, Rio de Janeiro, Brazil*, pages 23–36, 2014.

Nuel Belnap, Michael Perloff, and Ming Xu.

Facing the Future (Agents and Choices in Our Indeterminist World).
Oxford University Press, 2001.

M. Bratman.

Intentions, plans, and practical reasons.
Harvard University Press, 1986.

References II


Philip R. Cohen and Hector J. Levesque.
Intention is choice with commitment.
Artif. Intell., 42(2-3):213–261, 1990.


Arthur C. Danto.
Basic actions.
American Philosophical Quarterly, 2(2):141–148, 1965.


Donald Davidson.
The Logical Form of Action Sentences.
In N. Rescher, editor, *The Logic of Decision and Action*. University of Pittsburgh Press, 1967.


Donald Davidson.
Essays on Actions and Events.
Oxford University Press, 2001.


Barbara Dunin-Keplicz and Rineke Verbrugge.
Teamwork in Multi-Agent Systems - A Formal Approach.
Wiley series in agent technology. Wiley, 2010.


Davide Grossi, John-Jules Ch. Meyer, and Frank Dignum.
The many faces of counts-as: A formal analysis of constitutive rules.
J. Applied Logic, 6(2):192–217, 2008.


A. Goldman.
A Theory of Human Action.
Prentice-Hall, 1970.

References III

Andreas Herzig, Emiliano Lorini, and Nicolas Troquard.

A dynamic logic of institutional actions.

In *Computational Logic in Multi-Agent Systems - 12th International Workshop, CLIMA XII, Barcelona, Spain, July 17-18, 2011. Proceedings*, pages 295–311, 2011.

A. Kenny.

Action, Emotion, and Will.

Psychology Press, 1963.

Daniele Porello, Emanuele Bottazzi, and Roberta Ferrario.

The ontology of group agency.

In *Formal Ontology in Information Systems - Proceedings of the Eighth International Conference, FOIS 2014, September, 22-25, 2014, Rio de Janeiro, Brazil*, pages 183–196, 2014.

Daniele Porello.

Logics for modelling collective attitudes.

Fundam. Inform., 158(1-3):239–275, 2018.

Daniele Porello and Nicolas Troquard.

A resource-sensitive logic of agency.

In *Proceedings of the 21st European Conference on Artificial Intelligence (ECAI 2014)*, pages 723–728, August 2014.

Daniele Porello and Nicolas Troquard.

Non-normal modalities in variants of linear logic.

Journal of Applied Non-Classical Logics, 25(3):229–255, 2015.

A. S. Rao and M. P. Georgeff.

Modeling rational agents within a BDI-architecture.

In *Proceedings of the Second International Conference on Principles of Knowledge Representation and Reasoning*, pages 473–484. Morgan Kaufmann, 1991.

References IV

J. Searle.

Collective intentions and actions.

In P. Cohen, J. Morgan, and M.E. Pollack, editors, *Intentions in Communication*. MIT Press, 1990.

J. Searle.

The Construction of Social Reality.

The Free Press, 1995.

R. Tuomela and W. Balzer.

A fixed point approach to collective attitudes.

In G Holmström-Hintikka and R. Tuomela, editors, *Contemporary Action Theory*, volume II: Social Action, pages 115–142. Kluwer Academic Publishers, 1997.

R. Tuomela and K. Miller.

We-intentions.

Philosophical Studies, 53:367–389, 1988.

Raimo Tuomela.

The Importance of Us.

Stanford University Press, 1995.