Foundations for a BFO Conformant Ontology of Intentions and Plans

Alec Sculley

1. ***Introduction***

An ontology of intentions and plans is needed to model human cognition, and agency of both human agents and artificial intelligence agents.[[1]](#footnote-1) Existing BFO conformant ontologies that include the term ‘plan’ are the Common Core Ontologies (CCO),[[2]](#footnote-2) and the Industrial Ontologies Foundry (IOF),[[3]](#footnote-3) which give an informational account, and the Ontology for Biomedical Investigations (OBI),[[4]](#footnote-4) which gives a realizable entity account.

Remarkably, the existing ontologies do not both model intentions and plans despite using some cognate of ‘intention’ or related terms like ‘commitment.’

* Cco:Plan is a “Directive Information Content Entity that prescribes some set of **intended** **Intentional** Acts through which some Agent expects to achieve some Objective.”[[5]](#footnote-5)
* Iof:Plan is a “**intention**-to-perform processes on the part of an agent as prescribed by a plan specification.”[[6]](#footnote-6)
* Obi:Plan: is a “realizable entity that is the inheres in a bearer who is **committed** to realizing it as a planned process.”[[7]](#footnote-7)

Toyoshima, Barton, and Grenier give a dispositional account of intentions (and no account of plans),[[8]](#footnote-8) but, as will be shown, it falls short in important ways that a representational account does not. This paper ultimately gives a representational account of intentions and plans. It does so by using previous works on representations, namely, the Cognitive Process Ontology (CPO)[[9]](#footnote-9) and John Searle’s work on representations.[[10]](#footnote-10) This paper also uses Searle’s work on intentions,[[11]](#footnote-11) and Michael Bratman’s work on intentions and plans.[[12]](#footnote-12) After presenting the representational ontology of intentions and plans, this paper provides a formalization of the ontology, and discusses how the ontology relates to other works on intentions and plans.

1. ***Method: A (Meta)Representational Approach***

This paper uses the framework offered by Basic Formal Ontology (BFO), which asserts that every entity that exists is either a continuant or an occurrent.[[13]](#footnote-13) This paper also uses the Common Core Ontologies (CCO), which is a suite of mid-level ontologies that extends BFO toward domains of interest.[[14]](#footnote-14) Next, this paper uses the Cognitive Process Ontology (CPO) which extends CCO to cover cognitive processes and related entities, like representations and conative representations.[[15]](#footnote-15) To extend CPO to cover intentions, this paper uses the work of John Searle and Michael Bratman on intentions.[[16]](#footnote-16) To extend CPO to cover plans, this paper uses the work of Dan Sperber on metarepresentations.[[17]](#footnote-17)

Continuants are…

Independent continuants are

Generically dependent continuants are

Specifically dependent continuants are

Qualities are

Occurrents are…

Processes are…

CCO’s model of information uses three classes: information content entity, information bearing entity, and information quality entity.[[18]](#footnote-18) Information content entities are generically dependent continuants that are about some entity. As generically dependent continuants, information content entities generically depend on one or more independent continuant for their existence. For example, the information content entity that is Tolstoy’s *War and Peace*, which originally depended on the first completed manuscript of *War and Peace* but which now depends on every copy of *War and Peace* that exists. In particular, information content entities depend on some information bearing entity, which is an independent continuant that carries it. For example, the ink in this copy of *War and Peace*. Information content can only be carried by an independent continuant that bears some information quality entity, which is a quality that concretizes some information content entity. For example, the pattern of the ink in this copy of *War and Peace*.

This paper uses a version of CPO’s model of representations which is modified out of necessity. CPO defines a representation as “a quality that concretizes, or is intended to concretize, some information content entity.”[[19]](#footnote-19) The clause ‘or is intended to concretize’ is problematic for a project modelling intention because it seems to smuggle in some understanding of intention. While it is true that it must smuggle in some understanding of intention, the disjunctive nature of the clause does not strictly pose a problem for this project. However, since it is a defined class, it does indicate that there are two kinds of representations, and only one of them can be the basis for an ontology of intentions and plans. The first kind is just an information quality entity. The second is an intended representation, which causally depends on an intended act: an act prescribed by, and itself caused by, some intention. Clearly intended representations presuppose the existence of an intention and thus cannot be the ontological basis for intentions. I will now discuss representations generally, but will have in mind information quality entities that are not intended representations and thus omit the problematic clause. Later in the paper, after the representational model of intentions is presented, I will say more about intended representations.

A representation is a quality, a type of specifically dependent continuant, that concretizes some information content entity. For example (again), the pattern of the ink that concretizes this copy of *War and Peace*. As specifically dependent continuants, representations specifically depend on exactly one independent continuant for their existence. For example, this copy of *War and Peace* is borne by the ink on the page such that if the ink ceases to exist so does the pattern of the ink and thus this copy of *War and Peace*. As qualities, representations are realized at all times they exist, and require no extra process to be realized in contrast to realizable entities.

When representations concretize information content entities, they are derivatively about whatever the concretized information content entity is about.[[20]](#footnote-20) For example, this representation on the television screen is derivatively about World War Two because it concretizes some World War Two documentary, which is an information content entity about World War Two. This copy of *The History of the Peloponnesian War* is derivatively about the Peloponnesian War because it concretizes *The History of the Peloponnesian War*, which is an information content entity about the Peloponnesian War.

In addition to representations in books and on television screens, there are mental representations, which are the entities that humans (and other organisms) think with. CPO models mental representations using ‘cognitive system,’ and ‘mental quality.’ A cognitive system is a “system which realizes cognitive dispositions, all of whose parts are also parts of a single organism.”[[21]](#footnote-21) A mental quality is a “quality which specifically depends on an anatomical structure in the cognitive system of an organism.”[[22]](#footnote-22) A mental representation is a “mental quality that is a representation.”[[23]](#footnote-23)

Following Searle, both mental representations and non-mental representations, like maps, succeed or fail in matching reality.[[24]](#footnote-24) For example, my belief that it is snowing outside matches reality if and only if it is snowing outside, and fails to match reality if and only if it is not snowing outside. A map of New York City matches reality if and only if it is New York City pictured on the map. To account for this, Searle argues, we need two notions: success condition and direction of fit.[[25]](#footnote-25) The success condition is the way the world needs to be for the representation to match reality. The success condition in the given example is that it is snowing outside. The direction of fit of a representation specifies the direction of the matching relationship between the world and the representation.[[26]](#footnote-26) For example, beliefs have a representation-to-world direction of fit. That is to say that the success conditions of a belief can be satisfied only if it matches the way the world is. My belief that it is snowing, for example, is a successful belief if and only if it is snowing.

*Conative representations*, on the other hand, are what CPO calls a mental representation with a world-to-mind direction of fit.[[27]](#footnote-27) Examples include desires and, as will be shown, intentions.[[28]](#footnote-28) If I desire to eat pizza, my desire is satisfied if and only if I do eat pizza. Thus, the world must change to match my desire. Some conative representations concretize some descriptive information content entity that describes, for example, what one wants. These are plausibly desires.[[29]](#footnote-29) For example, some mental representation that concretizes some information content entity that describes that I want pizza can be satisfied if and only if I eat pizza. I can also want it to be snowing outside, and that representation’s success conditions is satisfied if and only if it does snow outside. In the first example, my representation must cause that success conditions (my eating pizza). In the second example, my representation cannot cause the satisfaction of the success conditions (this snowing outside).

Metarepresentations are representations that are derivatively about representations. Dan Sperber writes that there are four main types of metarepresentation:

Mental representations of mental representations (e.g., the thought “John believes that it will rain”), mental representations of public representations (e.g., the thought “John said that it will rain”), public representations of mental representations (e.g., the utterance “John believes that it will rain”), and public representations of public representations (e.g., the utterance “John said that it will rain”).[[30]](#footnote-30)

This paper will concern mental representations of mental representations, and public representations of mental representations.

1. ***(CODER)***

This section presents (CODER), the driving example used in this paper. Jim, a self-employed software engineer, begins his workday by deciding what he will do. His decision results in the intention to produce a program (objective) by writing one-hundred lines of code (action). Jim creates a plan to carry out his intention. First, he must represent his intention, either mentally or externally, such that it can be edited. Second, he must edit his objective intention such that it can more effectively constrain his action. So, he further specifies his objective. Third, he must edit his action intention such that it can more effectively guide him toward his more specific objective.

1. ***Informal Ontology***

This section presents the view forwarded by this paper in an informal manner. It does so by first addressing some theoretical foundations of the view, and then presenting the view. Not jumping straight into the formal ontology is intended to make the view more understandable.

* 1. ***Intentions***

Start with Searle and Bratman on intentions. First, following Searle, intentions are mental representations with a mind-to-world direction of fit. In order to successfully carry out some intention, the intending agent must make the world match the intention. Intentions, therefore, are conative representations. Second, following Bratman, intentions are “conduct-controlling pro-attitudes.”[[31]](#footnote-31) Intentions are pro-attitudes because they guide conduct including mental conduct like reasoning. Intentions are conduct controlling because they direct conduct, resist reconsideration, and are stable (Bratman calls this “inertia”).[[32]](#footnote-32) In other words, intentions are pro-attitudes to which an agent commits.

In (CODER), Jim starts the work day with the intention to produce a program (objective) by writing one-hundred lines of code (action). It does not matter whether Jim’s intention is reasonable. The program he intends to write might be impossible for one person to write in a day despite its relatively short length. This suggests that intentions are representations because Jim has it if and only if he has the correct attitude toward the processes he intends. The intention does not require any particular disposition that is realizable in the process he intends.

Suppose Jim’s intentions are reasonable. It is reasonable to think that Jim can produce the program by writing one-hundred lines of code. Further, Jim can carry out his intentions by producing the program by writing one-hundred lines of code. This makes his intention a conative representation because it has a world-to-mind direction of fit. However, in order to distinguish between intentions and other conative representations, we need to account for the attitude that makes intentions controlling. Following Bratman, this attitude is commitment.

Informally, then, we can define intention as follows.[[33]](#footnote-33)

* Intention is a conative representation that derivatively prescribes some process that one is committed to performing.

But Jim’s intention is what I will call a *compound intention*. That is, it is the intention to perform two or more processes, which are all related somehow. Following Searle, Jim intends to produce the program *by means of* writing one-hundred lines of code. The ‘by means of’ relation is causal: it is the writing one-hundred lines of code which (partially) causes the production of the program.[[34]](#footnote-34)

Informally, then, we can define causal compound intention as follows.

* Causal compound intention is an intention that derivatively prescribes some process that one is committed to performing by means of some other process.

Importantly, but, perhaps not necessarily,[[35]](#footnote-35) intentions result from some decision process, which can be both deliberative and non-deliberative. Decision processes are quite limited in the complexity of intention that they can output. They output representations, in the case of intentions, that are derivatively about things external to the deciding agent, in particular, processes.

* 1. ***Plans***

Start with Bratman on having a plan. Bratman argues that what it is to have a plan is to have a complex set of intentions. But, intentionally, Bratman does not discuss plans generally, only committing to a plan. In this paper, the goal is to model plans generally. Bratman says very little about plans generally, but as a starting point we can take what he does say. Plans generally, Bratman says, are “abstract structure[s] of a sort that can be represented, say, by some game-theoretical notation”[[36]](#footnote-36) as “some sort of partial function from circumstances to action”[[37]](#footnote-37) “that may be used to describe the [intentions] of different people.”[[38]](#footnote-38) So, plans generally, can be thought of as representations that describe the intentions of different people.

Now return to (CODER). In order for Jim to accomplish his original intention (i.e., to produce a program by means of writing one-hundred lines of code) he must turn it into a plan. This is because writing a program takes more deliberation about what to do than is involved in the decision process that outputs intentions. Creating a plan requires Jim to reason about his intentions, and to represent them in a way that can be edited and organized, typically, hierarchically. That is, creating a plan requires the planner to *meta*represent some intention. Each metarepresentation of some intention I will call a plan component. In other words, plan components are representations of intentions (which are representations). Defined informally,

* A plan component is a representation that derivatively represents some intention.

Plans are collections of plan components. Recipes, in some cases, are such that the order of actions listed do not need to be performed in order. So, defined informally,

* A plan is a group of plan components taken as a whole.

Returning to Jim, we might think that his plan must need to be organized hierarchically. For example, his plan might just begin as a two-item list which consists of his original single intention represented as two plan components.

(1) Write one-hundred lines of code.

(2) Produce program.

The order of execution matters here. Jim cannot produce the program without first writing the one-hundred lines of code. The second version of Jim’s plan might narrow the scope of his objective.

(1) Write one-hundred lines of code.

(2) Produce SPARQL program to answer the question ‘How many cities in the world have more than three professional sports teams?’

The third version of Jim’s plan might specify in more detail the actions Jim needs to take to achieve his objective.

(1) Write one-hundred lines of code.

(1.1.) Select cities and teams

(1.2.) Count teams per city

(1.3.) Filter cities with three or more teams

(1.4.) Count number of cities

(2) Produce SPARQL program to answer the question ‘How many cities in the world have more than three professional sports teams?’

Notably, (1.1.) – (1.4.) are representations of Jim’s intentions for his computer. Typically, people form intentions that derivatively prescribe processes that they will participate in. This, however, need not be the case. The same is true of plans.

1. ***Formal Ontology***

This section includes formalizations of each entity introduced in the previous section. It also includes the formalization of additional entities: (i) to expand the core beyond the driving example; and (ii) to show that the model is consistent with how theorists have thought about intentions and plans.

* 1. ***Core Formalization***

Begin with intention. To formalize it, this paper imports ‘action specification’ from CPO. An *action specification* is a “directive information content entity that prescribes some action.”[[39]](#footnote-39) This paper also uses a slightly modified version of ‘objective specification’ from CPO.[[40]](#footnote-40) An *objective specification* is a “directive information content entity that prescribes a process endstate.”[[41]](#footnote-41) This paper also imports the ‘is fused with’ object property from CPO (citing Husserl citing Stumpf). *Is fused with* is a primitive relation that relates two quality instances “so closely related that a new quality instance seems to emerge from them.”[[42]](#footnote-42) This paper also introduces two new terms. First, a *commitment value* is a mental quality that, when fused with a conative representation CR, determines the extent to which a cognitive system operates as if CR is controlling. Further, a *positive commitment value* is commitment value that when fused with a conative representation CR results in the bearing cognitive system treating CR, to some extent, as if it were controlling.[[43]](#footnote-43)

‘*x* is an intention’ means:

1. *x* is a conative representation

And there is some *y*, *z*, *w* such that

1. *y* is an action specification
2. *z* is an objective specification
3. *w* is a positive commitment value
4. *x* concretizes either
   1. *y*; or
   2. *z*; or
   3. Both *y* and *z*
5. *x* is fused with *w*

‘*x* is an action intention’ means:

1. *x* is an intention

And there is some *y* such that

1. *y* is an action specification
2. *x* concretizes *y*

‘*x* is an objective intention’ means:

1. X is an intention

And there is some *y* such that

1. *y* is an objective specification
2. X concretizes *y*

‘*x* is a compound intention’ means:

1. X is an intention

And there is some *y*, *z*, *w*, *u* such that

1. *Y* is some action specification
2. *Z* is some objective specification
3. *W* is some action
4. *U* is some objective
5. *Y* prescribes *w*
6. *Z* prescribes *u*
7. X concretizes y and z

‘*x* is a causal compound intention’ means:

1. X is an intention

And there is some *y*, *z*, *w*, *u* such that

1. *Y* is some action specification
2. *Z* is some objective specification
3. *W* is some action
4. *U* is some objective
5. *Y* prescribes *w*
6. *Z* prescribes *u*
7. *W* causes *u*
8. X concretizes y and z

‘*x* is a constitutive compound intention’ means:

1. X is an intention

And there is some *y*, *z*, *w*, *u* such that

1. *Y* is some action specification
2. *Z* is some objective specification
3. *W* is some action
4. *U* is some objective
5. *Y* prescribes *w*
6. *Z* prescribes *u*
7. *W* is occurrent part of *u*
8. X concretizes y and z

‘*x* is a plan component specification’ means:

1. *x* is a representational information content entity

And there is some *y* such that

1. *y* is an intention
2. *x* represents *y*

‘*x* is a plan component’ means:

1. *x* is a representation

and there is some *y* such that

1. *y* is a plan component specification
2. *x* concretizes *y*

‘*x* is a plan’ means:

1. *x* is a representation complex

and there is some *y* such that

1. *y* is a plan component
2. ***x has part only y***

‘*x* is a hierarchical plan component’ means:

1. *x* is a plan component

and there is some *y* such that

1. *y* is a hierarchy value
2. *x* is fused with *y*

‘*x* is a hierarchical plan’ means:

1. *x* is a plan

and there is some *y* such that

1. *y* is a hierarchical plan component
2. ***x has part only y***
   1. ***More formalizations*** 
      1. ***Motor, Proximal, and Distal Intentions***
3. ***Discussion***
   1. ***Views of intentions***
   2. ***Views of plans***
      1. ***Bratman***
      2. ***Mele***

1. [↑](#footnote-ref-1)
2. (CUBRC, Inc. [2017] 2022) [↑](#footnote-ref-2)
3. (Drobnjakovic et al. 2022) [↑](#footnote-ref-3)
4. (“Ontology for Biomedical Investigations” [2017] 2023) [↑](#footnote-ref-4)
5. (CUBRC, Inc. [2017] 2022) [↑](#footnote-ref-5)
6. (Smith et al., n.d.), 10. [↑](#footnote-ref-6)
7. (“Ontology for Biomedical Investigations” [2017] 2023) [↑](#footnote-ref-7)
8. (Toyoshima, Barton, and Grenier 2020) [↑](#footnote-ref-8)
9. (“Cognitive Process Ontology” 2023) [↑](#footnote-ref-9)
10. (Searle 2009; 1983; 2001) [↑](#footnote-ref-10)
11. (Searle 2009; 1983; 2001) [↑](#footnote-ref-11)
12. (Bratman 1999a; 1999b) [↑](#footnote-ref-12)
13. [↑](#footnote-ref-13)
14. (CUBRC, Inc. [2017] 2022) [↑](#footnote-ref-14)
15. (“Cognitive Process Ontology” 2023) [↑](#footnote-ref-15)
16. [↑](#footnote-ref-16)
17. (Sperber 2000) [↑](#footnote-ref-17)
18. (Donohue 2017; CUBRC, Inc. [2017] 2022) [↑](#footnote-ref-18)
19. (Limbaugh et al. 2021) [↑](#footnote-ref-19)
20. This is similar to the distinction between original and derivative concretizers found in (Limbaugh et al. 2021). [↑](#footnote-ref-20)
21. (Limbaugh et al. 2021, 12) [↑](#footnote-ref-21)
22. (Smith and Ceusters 2015) [↑](#footnote-ref-22)
23. (Smith and Ceusters 2015) [↑](#footnote-ref-23)
24. (Searle 2001, 36) [↑](#footnote-ref-24)
25. (Searle 2001, 37) [↑](#footnote-ref-25)
26. (Searle 1983, 73) says direction of fit is a primitive because it cannot be analyzed in terms of simpler entities. [↑](#footnote-ref-26)
27. (“Cognitive Process Ontology” 2023; Limbaugh et al. 2021) [↑](#footnote-ref-27)
28. (Searle 2001) [↑](#footnote-ref-28)
29. (Toyoshima, Barton, and Grenier 2020) [↑](#footnote-ref-29)
30. (Sperber 2000, 3) [↑](#footnote-ref-30)
31. (Bratman 1999b, 20, 22) [↑](#footnote-ref-31)
32. (Bratman 1999b, 22) [↑](#footnote-ref-32)
33. This paper presents a formal model in the following section. [↑](#footnote-ref-33)
34. The formal representation includes the second kind of compound intention: constitutive compound intention. [↑](#footnote-ref-34)
35. Mele thinks that some intentions are not the result of any decision. [↑](#footnote-ref-35)
36. (Bratman 1999b, 28) [↑](#footnote-ref-36)
37. (Bratman 1983, 271; 1984, 379) [↑](#footnote-ref-37)
38. (Bratman 1984, 379) [↑](#footnote-ref-38)
39. (“Cognitive Process Ontology” 2023) [↑](#footnote-ref-39)
40. CPO’s original definition, as with CPO’s definition of ‘representation,’ relies on some understanding of the word ‘intended.’ So, this paper will omit the word ‘intended’ for now, and later discuss intended objective specifications. [↑](#footnote-ref-40)
41. (“Cognitive Process Ontology” 2023) [↑](#footnote-ref-41)
42. (“Cognitive Process Ontology” 2023; Limbaugh et al. 2021) [↑](#footnote-ref-42)
43. This is a modified version of the definition of cpo:confidenceValue from (“Cognitive Process Ontology” 2023; Limbaugh et al. 2021) [↑](#footnote-ref-43)