Reasoning about Choice

Wiebe van der Hoek[1], Nicolas Troquard[2], and Michael Wooldridge[3]

[1] University of Liverpool, UK
[2] Institute of Cognitive Sciences and Technologies (ISTC-CNR), Italy
[3] University of Oxford, UK

Abstract. We present a logic for reasoning about choice. Choice CTL (C-CTL) extends the well-known branching-time temporal logic CTL with *choice modalities*, "$\Diamond$" and "$\Box$". An example C-CTL formula is $\Diamond \mathbf{AF} happy$, asserting that there exists a choice that will lead to happiness. C-CTL is related to both STIT logics and temporal cooperation logics such as ATL, but has a much simpler and (we argue) more intuitive syntax and semantics. After presenting the logic, we investigate the properties of the language. We characterise the complexity of the C-CTL model checking problem, investigate some validities, and propose multi-agent extensions to the logic.

1 Introduction

If we are interested in building autonomous agents, then we must surely be interested in the notion of *choice*. After all, an autonomous agent is essentially a system that is at liberty to make its own choices. It is not surprising, therefore, that choice features prominently in formal studies of action and agency. For example, logics of "seeing to it that" ("STIT") have been used to formalise the notion of an agent choosing to bring about some state of affairs [4,14,5,19,12,9]; cooperation logics study the notion of collective strategic choice [22,2]; and deontic logics try to isolate "acceptable" choices from "unacceptable" ones [32,20]. The formalisms cited above have shed much light on the notion of choice, and have contributed greatly to our understanding of the kinds of languages and semantics that might be used to capture choice. However, none of the above formalisms is without problems. For example, STIT logics are notoriously hard for humans to understand; cooperation logics tend to have rather complex semantics, and can also be hard for humans to understand, particularly when dealing with negated cooperation modalities; and deontic logics remain fraught with philosophical difficulties.

Our aim in the present paper is to develop a logic for reasoning about choices and their consequences that is much simpler, both syntactically and semantically, than the formalisms we cite above. We call the logic "Choice CTL" (C-CTL). As the name suggests, C-CTL is an extension of the well-known branching-time temporal logic CTL [18]. C-CTL extends CTL with *choice modalities*, "$\Diamond$" and "$\Box$". These modalities are used to express the properties of choices available to an agent. A formula $\Diamond\varphi$ asserts that the agent has a choice that will lead to φ being true, while $\Box\varphi$ means that no matter what choice the agent makes, φ will be true. The specific interpretation that we give to "choice" is that a choice represents a *constraint* on behaviour; crudely, the set of

C.I. Chesñevar et al. (Eds.): AT 2013, LNAI 8068, pp. 4–18, 2013.

choices available to an agent will be the set of all ways it can constrain its behaviour (subject to certain simple coherence constraints, that will be described later). It may seem strange to think of making choices as constraining behaviour, since intuitively, choices seem to be positive things ("I will do this..."), and constraints seem to be negative things ("I will not do this..."). In fact, we argue that this reading of choice is quite natural. For example, if I choose to work on my paper tonight, then I must exclude other choices (going to a party, going to the cinema, watching TV). If I choose to vote for one political party, then that excludes voting for another. If I choose to accept the job offer from Stanford, then I cannot accept the offer from MIT; and so on. In C-CTL, choice modalities can be combined with CTL operators. An example C-CTL formula is $\Diamond \mathbf{AF} happy$, asserting that there exists a choice for the agent that will be guaranteed to eventually lead to happiness[1]. C-CTL is related to both STIT logics and temporal cooperation logics such as ATL, but has a much simpler and (we argue) more intuitive syntax and semantics.

2 Logics of Action, Choice, and Agency

Von Wright is generally considered to be the pioneer in the contemporary philosophy of action. Over decades, he has developed an extended theory of action, treating the notion of agency as a modality. That is, agency is seen as an intensional notion instead of a mere referent in the language:

> It would not be right, I think 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. To act is, in a sense, to *interfere* with "the course of nature". [31, p.36]

The formalisation of choice that we develop in this paper will be largely consistent with this view of action.

While von Wright proposed a semi-formal semantics, we are looking for a logical framework that allows us to specify a system involving acting agents, and hence support the verification of such a system, for example via model checking [15]. Belnap *et al.*'s STIT theory [5] is a philosophical account of action that is equipped with a very rich and formal semantics. One key feature is that, (like Chellas [13] and von Kutschera [30] before), STIT theory is based in a *branching* model of time. The notions of a *history* and *history contingency* are central to the STIT view of agency. Belnap *et al.* illustrate this with the following quote:

> When Jones butters the toast, for example, the nature of his act, on this view, is to constrain the history to be realized so that it must lie among those in which he butters the toast. Of course, such an act still leaves room for a good deal of variation in the future course of events, and so cannot determine a unique history; but *it does rule out all those histories in which he does not* butter the toast. [5, p.33] (emphasis added)

[1] Alas, we do not necessarily know what that choice is.

Clearly, Belnap *et al.* see an act in similar ways as von Wright. Agency presupposes agent-related indeterminism and an action is an interference with the course of nature. The view of choice we present in this paper is founded directly on the idea that acting is ruling out possible histories.

Also in common with von Wright and Belnap *et al.*, we take the perspective that an action is a modal notion. This is to be opposed to the *ontological* treatment of action. To understand the ontological view, consider Davidson's famous example of the statement "John buttered the toast slowly with a knife". We might formalise this statement in predicate logic as follows:

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

where e is a variable denoting an event, John and $\mathit{the\ toast}$ are constants, and $\mathbf{butter}/3$, $\mathbf{slowly}/1$, and $\mathbf{with\ a\ knife}/1$ are predicates. Action sentences are then seen as denoting some logical combination of such relations.

Philosophers have developed an extensive literature in the ontology of action. In contrast to the modal view, and as we have just exemplified, it is usual to take an action to be a particular kind of event [16]. Also, it is assumed in linguistics that verbs denote events [28] which can be categorized, and some relationship can exist between each others [29]. What makes an entity an *acting* entity of an event is generally acknowledged to be the *intentionality* in action. To be the agent of an event, one has to make a rational decision governed by one's beliefs and desires [3,17]. Bratman [8] built upon this idea, and proposed that intentions operate like a filter over every action in order to select the actions that are desired and believed to be successful. For more on the subject, Bennett's [6] is an excellent monograph on the ontology of action. For an exploration of the middle ground between the modal view of agency and ontological view on actions see [25].

STIT logics represent probably the largest body of work on agentive action in the philosophy literature [4]. In the earliest account of STIT, an agent is said to see to it that φ if there has been a choice of his at a moment strictly in the past (the witness moment) such that (1) this choice made sure that φ would be true at this instant, and (2) there is a history that has been ruled out by this choice along which φ is false at this instant. From the point of view of modal logic, this semantics is of course rather complex, and much of the subsequent literature on the subject has been concerned with simplification. For more recent developments, we refer the reader to [19,24,12,21,10,9]. Although the work cited above clearly has philosophical value, we argue that the STIT framework remains rather opaque. Whilst the semantics of STIT has an undeniable explanatory power to the notion of agency in branching-time, it seems very difficult to model real world scenarios with it.

3 A Logic of Choice

Choice CTL (C-CTL) is based on the well-known branching time temporal logic CTL [18]. Recall that CTL allows one to express properties of branching-time temporal structures by combining *path quantifiers* **A** ("on all paths...") and **E** ("on some path...") with *tense modalities* **X** ("in the next state"), **F** ("eventually"), **G** ("always"), and **U** ("until"). For example, the formula $\mathbf{AG}\neg \mathit{fail}$ is a CTL formula expressing a *system invariant*: on all computations starting from now, at all states on the computation, the system will not enter a "fail" state. The formula $\mathbf{EF}\mathit{happy}$ expresses a *reachability* property: there is a possible computation of the system, on which eventually, I am happy. Note that in CTL, a temporal operator *must* be prefixed with a path quantifier.

C-CTL extends CTL with *choice modalities*: $\Diamond$ and $\Box$. A formula $\Diamond\varphi$ means "the agent has a choice such that if it makes this choice, φ will hold", while the formula $\Box\varphi$ means "whatever choice the agent makes, φ will hold". Notice that choice modalities are unary, and the argument to a choice modality can be a CTL formula, or indeed a formula containing choice modalities. So, for example, the formula $\Diamond\mathbf{AF}\mathit{happy}$ can be read as asserting that "the agent can make a choice that will eventually lead to happiness", while the formula $\Box\mathbf{AG}\mathit{poor}$ can be read as meaning that "no matter what choice the agent makes, it will always be poor". To be slightly more precise, a choice formula $\Diamond\varphi$ asserts that *the agent can constrain its behaviour in such a way that φ holds*, while $\Box\varphi$ means *no matter how the agent constrains its behaviour, φ will hold.*

Starting from a set Φ of Boolean variables, the syntax of C-CTL is defined by the following grammar:

$$\varphi ::= \top \mid p \mid \neg\varphi \mid \varphi \vee \varphi \mid \Diamond\varphi \mid \mathbf{EX}\varphi \mid \mathbf{E}(\varphi\mathbf{U}\varphi) \mid \mathbf{AX}\varphi \mid \mathbf{A}(\varphi\mathbf{U}\varphi)$$

where $p \in \Phi$. Formulas built from $\top$, p, negation and disjunction only are called *objective formulas*. Given the above operators, we can derive the remaining CTL temporal operators as follows:

$$\begin{array}{ll} \mathbf{AF}\varphi \equiv \mathbf{A}(\top\mathbf{U}\varphi) & \quad \mathbf{EF}\varphi \equiv \mathbf{E}(\top\mathbf{U}\varphi) \\ \mathbf{AG}\varphi \equiv \neg\mathbf{EF}\neg\varphi & \quad \mathbf{EG}\varphi \equiv \neg\mathbf{AF}\neg\varphi \end{array}$$

The box operator for choice is defined as a dual of the diamond choice operator in the conventional way: $\Box\varphi \equiv \neg\Diamond\neg\varphi$.

The scenarios we model consist of a single agent inhabiting an environment. The environment can be in any of a set S of possible states; states are assumed to be discrete, and to keep things simple, we assume S is finite and non-empty. In any given state, the agent is able to perform actions, which will cause a deterministic change in state. To capture these actions, we use a *transition relation*, $R \subseteq S \times S$. The intended interpretation is that if $(s, s') \in R$ then when the system is in state s, the agent can perform some action that will transform the system into state s'. Note that actions themselves are not explicitly present in the semantics; it is straightforward to add actions to our models, for example by labelling transitions (s, s') with the action that causes the transition. To express properties of the system, we assume a finite set $\Phi = \{p, q, \ldots\}$ of Boolean variables. A valuation function $V : S \to 2^{\Phi}$ tells us which Boolean variables are true in which states.

A *Kripke structure*, $\mathcal{K}$, is then a triple $\mathcal{K} = (S, R, V)$ where S is a non-empty finite set of *states*, $R \subseteq S \times S$ is a total[2] binary relation on S, which we refer to as the *transition relation*, and $V : S \to 2^{\Phi}$ labels each state with the set of Boolean variables true in that state.

[2] Totality here means that for every $s \in S$ there is a $t \in S$ such that $(s, t) \in R$.

A *path*, ρ, through a transition relation R, is an infinite sequence of states $\rho = (s_0, s_1, \ldots)$ such that $\forall u \in \mathbb{N}$, we have $(s_u, s_{u+1}) \in R$. If $u \in \mathbb{N}$, then we denote by $\rho[u]$ the element indexed by u in ρ (thus $\rho[0]$ denotes the first element, $\rho[1]$ the second, and so on). For a state s in a transition system $M = (S, R, V)$ we say that a path ρ is a s*-path* if $\rho[0] = s$. Let $paths_R(s)$ denote the set of s-paths over R.

We now define a binary *choice accessibility relation* "$\sqsupseteq$" over transition relations: $R \sqsupseteq R'$ will mean that "R' is a possible choice given transition relation R". Formally, where R and R' are transition relations over state set S, we write $R \sqsupseteq R'$ to mean that:

1. $R \supseteq R'$; and
2. R and R' are both total relations.

We will also write $R' \sqsubseteq R$ for $R \sqsupseteq R'$. Observe that the choice accessibility relation $\sqsupseteq$ is both reflexive and transitive.

The satisfaction relation "$\models$" for C-CTL is defined between *pointed structures* $\mathcal{K}, s$ (where $\mathcal{K} = (S, R, V)$ and $s \in S$) and C-CTL formulas, as follows:

$\mathcal{K}, s \models \Diamond\varphi$ iff $\exists R'$ such that s.t. $R \sqsupseteq R'$ and $(S, R', V), s \models \varphi$
$\mathcal{K}, s \models \mathbf{AX}\varphi$ iff $\forall \rho \in paths_R(s) : \mathcal{K}, \rho[1] \models \varphi$
$\mathcal{K}, s \models \mathbf{EX}\varphi$ iff $\exists \rho \in paths_R(s) : \mathcal{K}, \rho[1] \models \varphi$
$\mathcal{K}, s \models \mathbf{A}(\varphi\mathbf{U}\psi)$ iff $\forall \rho \in paths_R(s), \exists u \in \mathbb{N}$, s.t. $\mathcal{K}, \rho[u] \models \psi$ and $\forall v, (0 \leq v < u) : \mathcal{K}, \rho[v] \models \varphi$
$\mathcal{K}, s \models \mathbf{E}(\varphi\mathbf{U}\psi)$ iff $\exists \rho \in paths_R(s), \exists u \in \mathbb{N}$, s.t. $\mathcal{K}, \rho[u] \models \psi$ and $\forall v, (0 \leq v < u) : \mathcal{K}, \rho[v] \models \varphi$

and in a standard way for the propositional connectives. As usual, we write $\models \varphi$ to indicate that $\mathcal{K}, s \models \varphi$ for all pointed structures $\mathcal{K}, s$.

Example 1. Consider the Kripke structure $\mathcal{K}$ displayed in Figure 1. In state s_1, the agent is at home (the atom h is true at s_1). When going to work, our agent has three options: he can first pick up a colleague (c) or his boss (b) before setting for the office, or he can stop at a coffee shop with time for a warm chocolate first. In the latter case, he would be selfish (s), in the first two cases, much more altruistic (a). Once work is finished, there are two options: returning home or, alternatively, retire (r).

The following statements are true in $\mathcal{K}, s_1$. Firstly, $\Diamond(\mathbf{AF}a \wedge \mathbf{EG}\neg b)$: by leaving the transition (s_1, s_4) out of the system, on all remaining paths the agent will eventually be altruistic, without having to ever take his boss to work. Similarly, we have $\Diamond(\mathbf{AF}a \wedge \Diamond\mathbf{AF}c)$: the agent can commit himself to always be altruistic, even in such a way that he can later on commit himself further to always take his colleague. Finally note that we have $\Diamond\mathbf{A}(\mathbf{G}\neg r \vee (h \vee a \vee w)\mathbf{U}r)$: there is a choice for the agent, so that in all remaining branches, he either will never retire, or else he will always be either home, or altruistic or at work until he retires.

Note, in $\mathcal{K}, s_1$ that it is unavoidable that the agent at least once goes to work: $\mathbf{AF}w$. One would expect that what is unavoidable is also true no matter which choice the agent makes, and indeed we have $\Box\mathbf{AF}w$. However, the formula $\mathbf{AF}\varphi \rightarrow \Box\mathbf{AF}\varphi$ is not a validity as the following counterexample $\varphi = \mathbf{EX}r$ demonstrates in our model: in s_1, it is true that in all paths the agent has the choice to retire and 'transit' to state s_6, (this

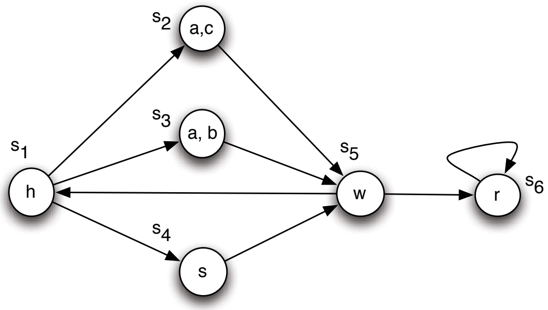

Fig. 1. A simple work-office example

is true even for the path $s_1, s_3, s_5, s_1, s_3, s_5, s_1, \ldots$), but it is not the case that, no matter which choice the agent makes, $\mathbf{EX}r$ is true (the agent can choose a transition relation that does not include (s_5, s_6)).

In a similar way, the formula $\Diamond\mathbf{EF}\varphi \rightarrow \mathbf{EF}\varphi$ is not a validity: it is not necessarily the case that if there is a choice that guarantees that along some path, some property φ is eventually true, then there should be a path where φ is eventually the case. Take φ to be $\mathbf{AX}a$. Indeed, in state s_1, the agent can make a choice (leave out the transition (s_1, s_4)) with the effect that on some path (like s_1, s_3, s_5, s_1), at some point (s_1), in the next state the agent is bound to be altruistic, i.e., $M, s_1 \models \Diamond\mathbf{EFAX}a$. However, we also have $M, s_1 \models \neg\mathbf{EFAX}a$: it is not the case that there is a path such that at some point along it, the agent is bound to be altruistic: the agent has not committed himself to anything yet!

The truth value of $\Diamond$ allows the agent to restrict the current relation to any total sub-relation. It might seem extreme at first sight to allow an agent to be able to restrict the relation R to a mere function, hence associating deterministically a state to its successor. This is however in perfect compatibility with the assumption that agents are the source of indeterminism. And typically, agents can plan ahead for any state of the game. These are for instance the common assumptions in game theory, where the future is completely determined when all agents have made their choice. Our proposal is consistent with this view: we will see in Section 5 that the multi-agent variation of our logic satisfies the most common properties of logics for games and for social choice theory.

Before proceeding, we will consider the model checking problem for C-CTL [15].

MODEL CHECKING:
Instance: Kripke structure $\mathcal{K} = (S, R, V)$, state $s \in S$, and C-CTL formula φ.
Question: Is it the case that $\mathcal{K}, s \models \varphi$?

The model checking problem for the underlying temporal logic CTL is P-complete [23]; however, adding choice modalities to the language complicates the decision problem considerably:

Theorem 1. *The* MODEL CHECKING *problem is PSPACE-complete.*

4 Towards an Axiomatization

We are not in a position to offer a complete axiomatization of C-CTL at this point, and the aim of this section is instead to show a number of validities which are candidates for axioms. In addition, we point out some interesting properties of our language, which give some indication of why a complete axiomatization is not easy to obtain.

By our observation that the choice relation $\sqsupseteq$ is reflexive and transitive, we immediately get the following:

$$\models \varphi \to \Diamond\varphi \quad \text{and} \quad \models \Diamond\Diamond\varphi \to \Diamond\varphi \tag{1}$$

The first validity of (1) expresses that if something is true in the current system based on the transition relation R, the agent can make a choice (namely, R), such that φ. The second validity expresses that a restriction of a restriction of R is a restriction of R. The "dual" of the first validity of (1) is $\models \Box\varphi \to \varphi$, i.e., the modal scheme T. The converse of this scheme is obviously not true for all φ, but it *does* hold for purely propositional formulae:

$$\models \psi \to \Box\psi \qquad \text{if } \psi \text{ is propositional} \tag{2}$$

As a property of choice, (2) makes perfect sense: choice interferes with the future, but no choice can change the actual facts.

The properties of (1) make the operator $\Diamond$ a $KT4$ operator [7], however, the fact that we want (2) for objective formulas φ but not for *arbitrary* formulas, implies that an axiomatization for C-CTL would not include the principle of *uniform substitution*: to derive (2) for objective formulas, we would add the axioms $p \to \Box p$ and $\neg p \to \Box\neg p$ (*atomic permanence*) for $p \in \Phi$, and not for arbitrary φ. (In fact, atomic permanence follows from yet another axiom that we will discuss, and which only involves atoms, $Ax7$).

How about the relation between the agent's choices and possible futures, i.e., the relation between $\Diamond$ on the one hand, and $\mathbf{E}$ and $\mathbf{A}$ formulas on the other hand? As we argued in Example 1, the formula $\Diamond\mathbf{EF}\varphi \to \mathbf{EF}\varphi$ is not a validity, the counterexample for φ being an $\mathbf{A}$ formula. However, an example of a related formula that *is* valid is:

$$\models \Diamond\mathbf{EFEG}p \to \mathbf{EFEG}p \tag{3}$$

The validity in (3) expresses that if the agent can make a choice such that, as a consequence, there is a path such that at some time there is a path such that p is true along the path, then there is a path where that consequence is already true.

The validity above reflects a property in first-order logic that universal formulas are preserved when taking submodels and existential formulas are preserved under taking supermodels. To formalise the validities we are after, we define two sublanguages of C-CTL: the universal language L^u (with typical element μ), and the existential fragment L^e (typical element ε):

$$\mu ::= \top \mid \bot \mid p \mid \neg p \mid \mu \vee \mu \mid \mu \wedge \mu \mid \mathbf{AX}\mu \mid \mathbf{AG}\mu \mid \mathbf{A}(\mu\mathbf{U}\mu) \mid \Box\mu$$
$$\varepsilon ::= \top \mid \bot \mid p \mid \neg p \mid \varepsilon \vee \varepsilon \mid \varepsilon \wedge \varepsilon \mid \mathbf{EX}\varepsilon \mid \mathbf{EG}\varepsilon \mid \mathbf{E}(\varepsilon\mathbf{U}\varepsilon) \mid \Diamond\epsilon$$

The following theorem is a generalisation of [26].

Theorem 2. *We have the following:*

$$\forall \varepsilon \in L^e : \models \Diamond\varepsilon \to \varepsilon \quad \textit{and}$$
$$\forall \mu \in L^u : \models \mu \to \Box\mu.$$

Note that both (2) and (3) are instances of Theorem 2.

There are validities not captured by Theorem 2. Consider a modal logic with two diamonds $\Diamond_1$ and $\Diamond_2$, each associated with an accessibility relation R_1 and R_2, respectively. Now consider the scheme $\Diamond_1\Diamond_2 p \to \Diamond_2\Diamond_1 p$. Semantically this expresses a grid-like property

$$\forall xyz((x,y) \in R_1 \ \& \ (y,z) \in R_2 \Rightarrow \exists v(x,v) \in R_2 \ \& \ (v,z) \in R_1) \tag{4}$$

Now, consider the two models $M = (W, R, V)$ and $M' = (W, R', V)$ with $R \sqsupseteq R'$ in Figure 2. The white-headed arrows denote a transition signalling that we are going to interpret path-quantifiers with respect to R', rather than to R. Call this relation R_1. Moreover, let R_2 be the relation that specifies the path $x, x_1, x_2, x_3, \ldots$: this is a path present in both M and M'. Then it is immediately clear from Figure 2 that we have the grid-like property 4. This then gives rise to the following, where φ is an arbitrary formula:

$$\models \Diamond\mathbf{EX}\varphi \to \mathbf{EX}\Diamond\varphi$$
$$\not\models \Diamond\mathbf{EG}\varphi \to \mathbf{EG}\Diamond\varphi \tag{5}$$

A counterexample for the non-validity is obtained by $\varphi = \mathbf{AX}p$ and with R' being a restriction of R that such that φ currently holds under that restriction, but not under R itself (i.e., there should be some path for which $\mathbf{X}p$ is true, and some for which $\mathbf{X}\neg p$).

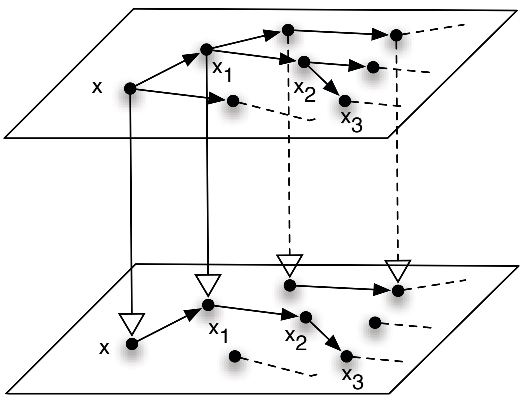

Fig. 2. A model $M = (S, R, V)$ (top) and $M' = (S, R', V)$ with $R \sqsupseteq R'$

Another validity of C-CTL is the following

$$\models \mathbf{EX}p \rightarrow \Diamond\mathbf{AX}p \tag{6}$$

The proof of (6) is simple: if we have $(W, R, V), s \models \mathbf{EX}p$ it means that there is an s-path $s, s_1, s_2, \ldots$ so that $(W, R, V), s_1 \models p$. Let R' be obtained from R by removing all transitions (s, t) from R for which $t \neq s_1$: this has as an effect that all paths based on R' (which is still a total relation) from s have to go through s_1, hence, in all of those paths, $\mathbf{E}p$ is true.

We already know that uniform substitution is not valid for C-CTL, so we cannot expect $\mathbf{EX}\varphi \rightarrow \Diamond\mathbf{AX}\varphi$ to be true for arbitrary φ. It is not difficult though to see that the choice $\varphi = \mathbf{EX}p$ gives another validity: $\mathbf{EXEX}p \rightarrow \Diamond\mathbf{AXEX}p$.

But given those validities, it may come as a surprise that we have

$$\not\models \mathbf{EXEXEX}p \rightarrow \Diamond\mathbf{AXEXEX}p$$

A simple demonstration of the displayed non-validity is provided in Figure 3: here, in $(S, R, V), s$ it holds that $\mathbf{EXEXEX}p$ (u is the only state where p is true). The argument that demonstrates this, uses the states s, t and u. Now, suppose we would have $\Diamond\mathbf{AXEXEX}p$ in s, then for some total subrelation R' of R, we have $(S, R', V), s \models \mathbf{AXEXEX}p$. We cannot take $R' = R$, since $\mathbf{AXEXEX}p$ is not true in $(S, R, V), s$. If R' is such that $(s, t) \notin R'$, then $\mathbf{AXEXEX}\varphi$ is only true for those φ that are true in v, which excludes p. If $(s, u) \notin R'$, then no p state is reachable any longer, so $\mathbf{AXEXEX}p$ does not hold in $(W, R', V), s$. This demonstrates that for no choice R' the formula $\mathbf{AXEXEX}p$ is true in s, so $(W, R, V), s \models \neg\Diamond\mathbf{AXEXEX}p$. Loosely formulated: in order to make $\Diamond\mathbf{AXEXEX}p$ true in s, we need the transition (s, u) to reach a p-state, but since paths through u continue only into $\neg p$-states, we would also like to to get rid of the transition (s, u).

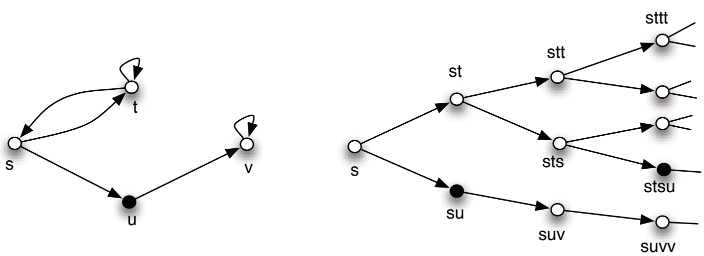

Fig. 3. A model M and its unraveling M' (atom p is true in black states)

This brings an interesting aspect of our language to light. Call a model $M = (W, R, V)$ *tree-like* if R represents a tree. Then, it is easy to see that in such models, we have (6) for *arbitrary* formulas φ, rather than p, since removing a transition from s has no repercussions for successors of s. Given an arbitrary model $M = (W, R, V)$, one can define its unraveling $M^u = (W^u, R^u, V^u)$, a tree-like model in which all possible paths from M are "unraveled". Rather than giving the formal definition (see [7, p. 63]), we refer to the model M' at the right hand side of 3, which is the unraveling of the model M on the left.

Observation 1. *Let M be a model and M' its unraveling.*

1. *For all $\Box$-free formulas: $M, w \models \varphi$ iff $M', s \models \varphi$*
2. *Let M and M' be the models of Figure 3, and let φ be $\Diamond\mathbf{AXEXEX}p$. Then $M, s \models \neg\varphi$, while $M', s \models \varphi$.*

Since a model M and its unraveling M' are a special case of models that are *bisimilar*, we have an argument for the non-modal behaviour of C-CTL, since modal languages are invariant under bisimulations.

In fact, we can define a bisimulation-like notion for C-CTL:

Definition 1. *Let $\mathcal{K}_1 = (S_1, R_1, V_1)$ and $\mathcal{K}_2 = (S_2, R_2, V_2)$ be two Kripke structures, $s_1 \in S_1$ and $s_2 \in S_2$. We say that $\mathcal{K}_1, s_1$ and $\mathcal{K}_2, s_2$ match, written $\mathcal{K}_1, s_2 \sim \mathcal{K}_2, s_2$, if the following holds:*

1. *$V_1(s_1) = V_2(s_2)$ (atomicity)*
2. *$\forall t_1 \in S_1((s_1, t_1) \in R_1 \Rightarrow \exists t_2 \in S_2$ such that $(s_2, t_2) \in R_2$ and $\mathcal{K}_1, t_1 \sim \mathcal{K}_2, t_2)$ (CTL-forth)*
3. *$\forall t_2 \in S_2((s_2, t_2) \in R_2 \Rightarrow \exists t_1 \in S_1$ such that $(s_1, t_1) \in R_1$ and $\mathcal{K}_1, t_1 \sim \mathcal{K}_2, t_2)$ (CTL-back)*
4. *For every $R_1' \sqsubseteq R_1$ there is $R_2' \sqsubseteq R_2$ such that $(W_1, R_1', V_1), s_1 \sim (W_2, R_2', V_2), s_2$ (Δ-forth)*
5. *For every $R_2' \sqsubseteq R_2$ there is $R_1' \sqsubseteq R_1$ such that $(W_1, R_1', V_1), s_1 \sim (W_2, R_2', V_2), s_2$ (Δ-back)*

Figure 4 shows examples of three matching models $\mathcal{K}_i, s_i$. We have:

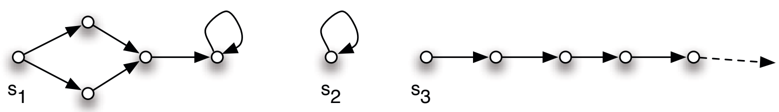

Fig. 4. Three matching models, assuming that all valuations agree

Lemma 1. *Suppose $\mathcal{K}_1, s_1 \sim \mathcal{K}_2, s_2$ match. Then they agree on all C-CTL formulas.*

In summary, an axiomatization for C-CTL (see Table 1) would need at least the axioms of CTL [18], axioms $Ax2$–$Ax4$, which regulate the behaviour of the modality $\Box$, and some mix axioms $Ax5$–$Ax7$.

Table 1. Some axioms for C-CTL

<u>CTL Axioms:</u>	
$(Ax1)$ CTL tautologies	
<u>Choice Axioms:</u>	
$(Ax2)$ $\Box(\varphi \rightarrow \psi) \rightarrow (\Box\varphi \rightarrow \Box\psi)$	
$(Ax3)$ $\Box\varphi \rightarrow \varphi$	
$(Ax4)$ $\Box\varphi \rightarrow \Box\Box\varphi$	
<u>Mix Axioms:</u>	
$(Ax5)$ $\mu \rightarrow \Box\mu$	$\mu \in L^u$
$(Ax6)$ $\Diamond\mathbf{EX}\varphi \rightarrow \mathbf{EX}\Diamond\varphi$	
$(Ax7)$ $\mathbf{EX}p \rightarrow \Diamond\mathbf{AX}p$	$p \in \Phi$
<u>Inference Rules:</u>	
$(IR1)$ From $\vdash \varphi$ and $\vdash \varphi \rightarrow \psi$ infer $\vdash \psi$	
$(IR2)$ From $\vdash \varphi$ infer $\vdash \mathbf{AG}\varphi$	
$(IR3)$ From $\vdash \varphi$ infer $\vdash \Box\varphi$	

5 Multi-agent Extensions

In this section, we show how C-CTL can be extended in a rather straightforward manner to capture the choice of coalitions of agents. We should emphasise that we are not concerned with agent's *motivations* for action. We did not take into account the agent's intention in the logic of the individual choice of Section 3, and we are not going to consider team attitudes in the multi-agent setting (see e.g., [33]). However, there is one aspect of collective agency that we aspire to. It is the aspect of power of coalitions that comes from social choice theory [1]. It has been translated in terms of modal logics with coalition logic [22], in Alternating-time Temporal Logic [2], and by extension in the STIT frameworks that embed them [12,11].

C-CTL can naturally be extended to multi-agent settings. We assume a system is populated by a set $Ag = \{1, \ldots, n\}$ of *agents*, and that the actions available to each agent $i \in Ag$ are captured by an individual transition relation $R_i \subseteq S \times S$. We refer to a collection of transition relations $R_1, \ldots, R_n$ (where there is a transition relation for each agent $i \in Ag$) as a *collective transition relation*. A Kripke structure is now defined to be a tuple $\mathcal{K} = (S, R_1, \ldots, R_n, V)$ where S and V are as defined before, and $R_1, \ldots, R_n$ is a collective transition relation. We extend the relation $\sqsupseteq$ defined earlier for individual transition relations to *coalitions*, C, which are simply subsets of agents $C \subseteq Ag$. We write

$$(R_1, \ldots, R_n) \sqsupseteq_C (R'_1, \ldots, R'_n)$$

to mean that:

1. $\forall i \in C$ we have $R_i \sqsupseteq R'_i$; and
2. $\forall j \in Ag \setminus C$ we have $R_j = R'_j$.

Given this definition, we can present the semantics of *Multi-agent* C-CTL (MC-CTL), as follows – note that the rules defining the propositional connectives and the path quantifiers remain unchanged, and we will not restate them.

$\mathcal{K}, s \models \langle C\rangle\varphi$ iff $\exists(R'_1, \ldots, R'_n)$ such that s.t. $(R_1, \ldots, R_n) \sqsupseteq_C (R'_1, \ldots, R'_n)$ and $(S, (R'_1, \ldots, R'_n), V), s \models \varphi$

We define the dual $[C]$ of the collective choice modality in the standard way:

$$[C]\varphi \equiv \neg\langle C\rangle\neg\varphi.$$

There is a close connection between our operator of choice and the notion of brute choice captured by the Chellas STIT. An agent sees to it that φ if given his current choice, φ is true whatever the other agents do. A modality similar to the Chellas STIT is then:

$$[C\ stit]\varphi \equiv [Ag \setminus C]\varphi$$

It is also straightforward to see that the $\sqsubseteq_C$ relation is reflexive and transitive, and so collective choice modalities satisfy the modal axioms K, T, and 4, as with individual choice. We can define ATL-like cooperation modalities as abbreviations, as follows:

$$\langle\langle C\rangle\rangle\varphi \equiv \langle C\rangle[C\ stit]\varphi$$

That is, C has the power to achieve φ, if there is a choice of C such that C sees to it that φ. These constructions are not new and have been already used for example in [27].

We argue that this operator does indeed behave very much like the cooperation modality in ATL/Coalition Logic CL. Table 2 shows some theorems of MC-CTL, which are direct counterparts of CL axioms (see, e.g., [22, p. 54]).

They are the syntactic representation of some core principles of social choice theory that regulate the powers of coalitions:

1. coalitions always have the power to achieve something;
2. if a coalition C_1 has the power to achieve φ, then every super-coalition $C_2 \supseteq C_1$ has the power to achieve φ;
3. if a coalition C_1 has the power to achieve φ and an independent coalition C_2 has the power to achieve ψ, then C_1 and C_2 have together the power to achieve $\varphi \wedge \psi$.

Table 2. Pauly's cooperation axioms hold for MC-CTL

CL Axioms:	
$(MCCTL1)$ $\neg\langle\langle Ag\rangle\rangle\bot$	
$(MCCTL2)$ $\langle\langle C_1\rangle\rangle\varphi \rightarrow \langle\langle C_2\rangle\rangle\varphi$	where $C_1 \subseteq C_2$
$(MCCTL3)$ $\langle\langle C_1\rangle\rangle\varphi \wedge \langle\langle C_2\rangle\rangle\psi \rightarrow \langle\langle C_1 \cup C_2\rangle\rangle(\varphi \wedge \psi)$	where $C_1 \cap C_2 = \emptyset$

Let us consider a multi-agent example.

Example 2. Consider the system depicted in Figure 5. We have three agents $Ag = \{1, 2, 3\}$. We call 1 and 2 the clients, and 3 the server. A resource is moved along: in s, agent 3 can either pass the resource to 1 (leading to state t) or to 2 (state x). For i and j agents, an arrow labelled $i : j$ denotes that i passes on the resource to j. If there is no outgoing edge from a state for agent i, we assume his only option is to wait, i.e., we have not drawn reflexive arrows labelled $i : w$. In state t where 1 has the resource, he can pass it back to 3 or he can choose to use it: the edge $1 : u$ denotes a transition from t to u. Similarly for agent 2 in state x. To reason about this scenario, we use atoms h_i (agent i holds the resource), u_i (agent i is using the resource) and b_i (agent i has benefited from the resource in the current cycle). Atom h_i is true in any state with an incoming arrow labelled $j : i$ for some j, atom u_i holds iff there is an incoming arrow labelled $i : u$, and b_1 is true in u, v, w and q, while b_2 is true in w, y, z, q. The black filled states satisfy $b_1 \wedge b_2$.

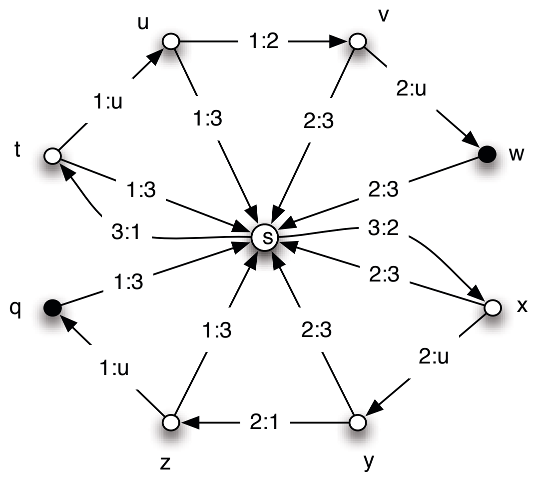

Fig. 5. A simple multi-agent C-CTL system

Assume the starting state is s. In $\mathcal{K}, s$ the following holds. First of all, the formula $\langle\langle 1, 2, 3\rangle\rangle \mathbf{AGAF}(b_1 \wedge b_2)$ does *not* hold: the grand coalition cannot constrain their choices in such a way that both clients are guaranteed they will benefit infinitely often from the resources. This is so because if no matter whether the system transits from s to t or to x, there will be an agent (2 at t and 1 at x) that has no choice but to generate a path $t, t, t, \ldots$ or $x, x, x, \ldots$, respectively. This shows that when we use the $\mathbf{A}$ quantifier, we quantify over all paths that nature can possibly choose, given the current constraints. It is easy to see that we do have: $\langle\langle 1, 2, 3\rangle\rangle \mathbf{EGEF}(b_1 \wedge b_2)$. We also have the following: $\langle\langle 1, 3\rangle\rangle \mathbf{EF}\langle\langle 2\rangle\rangle \mathbf{A}(h_1 \mathbf{U} h_2)$. That is, 1 and 3 together can constrain themselves in such a way that on some resulting path at some time it holds that 2 has a choice such that 1 holds the resource until 2 holds it. Note that agents have power to exclude each other from the resource: $\langle\langle 1, 3\rangle\rangle \mathbf{AG}\neg h_2$: agents 1 and 3 can constrain their actions in such a way that 2 never holds the resource. What can 3 choose? He cannot on his own prevent a client to hold the resource in the future, but he can determine the order in which they receive it: $\langle\langle 3\rangle\rangle \mathbf{AG}(h_2 \rightarrow b_1)$: the server 3 can organise his actions in such a way, that in all resulting computations it holds that whenever 2 holds the resource, 1 has already benefited from it. Also note that an agent $i = 1, 2$ can avoid using the resource, but cannot avoid holding it: $\langle\langle i\rangle\rangle \mathbf{AG}\neg u_i \wedge \neg\langle\langle i\rangle\rangle \mathbf{EG}\neg h_i$.

6 Conclusions

We have grounded in the philosophy of action the idea that choosing is to rule out some courses of nature. We have then proceeded to present a logic C-CTL with one agent that follows this idea, and considered multi-agent extensions. We have seen that the notion of powers of agents and coalitions that it reflects is consistent with the theories one can find in social choice theory. In contrast with the other logics of choice in the literature based on branching-time models, our examples demonstrate the ease with which our logic makes it possible to model quite complex systems of interacting agents.

One obvious development of this work would be a complete axiomatization, and to characterise the complexity of the satisfiability problem. The connection to game theoretic reasoning could be explored, and, related to this, an important extension of C-CTL might be obtained by having ways to reason about the knowledge that agents have about their choices, and that of others.

Acknowledgements. Wooldridge was supported by the European Research Council under Advanced Grant 291528 (“RACE”). Troquard was supported by a fellowship Marie Curie Action People Trentino.

References

1. Abdou, J., Keiding, H.: Effectivity Functions in Social Choice Theory. Kluwer Academic Publishers, Dordrecht (1991)
2. Alur, R., Henzinger, T.A., Kupferman, O.: Alternating-time temporal logic. Journal of the ACM 49(5), 672–713 (2002)
3. Anscombe, G.E.M.: Intention, 2nd edn. Blackwell (1963)
4. Belnap, N., Perloff, M.: Seeing to it that: a canonical form for agentives. Theoria 54, 175–199 (1988)
5. Belnap, N., Perloff, M., Xu, M.: Facing the future: agents and choices in our indeterminist world, Oxford (2001)
6. Bennett, J.: Events and their names. Hackett Publishing Company, Indianapolis (1988)
7. Blackburn, P., de Rijke, M., Venema, Y.: Modal Logic. Cambridge University Press, Cambridge (2001)
8. Bratman, M.: Intentions, plans, and practical reason. Harvard University Press, Cambridge (1987)
9. Broersen, J.: CTL.STIT: enhancing ATL to express important multi-agent system verification properties. In: Proceedings of AAMAS 2010 (2010)
10. Broersen, J., Herzig, A., Troquard, N.: A STIT-extension of ATL. In: Fisher, M., van der Hoek, W., Konev, B., Lisitsa, A. (eds.) JELIA 2006. LNCS (LNAI), vol. 4160, pp. 69–81. Springer, Heidelberg (2006)

11. Broersen, J., Herzig, A., Troquard, N.: Embedding Alternating-time Temporal Logic in Strategic STIT Logic of Agency. Journal of Logic and Computation 16(5), 559–578 (2006)
12. Broersen, J., Herzig, A., Troquard, N.: Normal coalition logic and its conformant extension. In: Samet, D. (ed.) Proceedings of the Eleventh Conference Theoretical Aspects of Rationality and Knowledge (TARK), pp. 42–51. Presses universitaires de Louvain, Brussels (2007)
13. Chellas, B.: The Logical Form of Imperatives. Perry Lane Press, Stanford (1969)
14. Chellas, B.F.: On bringing it about. Journal of Philosophical Logic 24(6), 563–571 (1995)
15. Clarke, E.M., Grumberg, O., Peled, D.A.: Model Checking. The MIT Press, Cambridge (2000)
16. Davidson, D.: The Logical Form of Action Sentences. In: Rescher, N. (ed.) The Logic of Decision and Action, pp. 81–120. University of Pittsburgh Press (1967)
17. Dennett, D.C.: Intentional systems. Journal of Philosophy, 68(4) (1971)
18. Emerson, E.A.: Temporal and modal logic. In: van Leeuwen, J. (ed.) Handbook of Theoretical Computer Science Volume B: Formal Models and Semantics, pp. 996–1072. Elsevier Science Publishers B.V, Amsterdam (1990)
19. Horty, J.: Agency and Deontic Logic. Oxford University Press, Oxford (2001)
20. Meyer, J.-J.C., Wieringa, R.J. (eds.): Deontic Logic in Computer Science — Normative System Specification. John Wiley & Sons (1993)
21. Müller, T.: On the formal structure of continuous action. In: Schmidt, R., Pratt-Hartmann, I., Reynolds, M. (eds.) AiML-2004: Advances in Modal Logic, pp. 277–286 (2004)
22. Pauly, M.: Logic for Social Software. PhD thesis, University of Amsterdam, ILLC Dissertation Series 2001-10 (2001)
23. Schnoebelen, P.: The complexity of temporal logic model checking. In: Balbiani, P., Suzuki, N.-Y., Wolter, F., Zakharyascev, M. (eds.) Advanced in Modal Logic, vol. 4, pp. 393–436. King's College Publications, London (2003)
24. Troquard, N.: Independent agents in branching-time. PhD thesis, Univ. of Toulouse & Univ. of Trento (2007)
25. Troquard, N., Trypuz, R., Vieu, L.: Towards an ontology of agency and action: From STIT to OntoSTIT+. In: Bennett, B., Fellbaum, C. (eds.) International Conference on Formal Ontology in Information Systems (FOIS), Baltimore, Maryland, USA. Frontiers in Artificial Intelligence and Applications, vol. 150, pp. 179–190. IOS Press,
26. van der Hoek, W., Roberts, M., Wooldridge, M.: Social laws in alternating time: Effectiveness, feasibility, and synthesis. Synthese 156(1), 1–19 (2007)
27. van der Hoek, W., Wooldridge, M.: On the logic of cooperation and propositional control. Artificial Intelligence 64, 81–119 (2005)
28. Vendler, Z.: Verbs and times. Philosophical Review 66, 143–160 (1957)
29. Verkuyl, H.: A Theory of Aspectuality. Cambridge University Press (1993)
30. von Kutschera, F.: Bewirken. Erkenntnis 24(3), 253–281 (1986)
31. von Wright, G.H.: Norm and Action: A Logical Inquiry. Routledge & Kegan Paul, London (1963)
32. von Wright, G.H.: Deontic logic. Mind 60(237), 1–15 (1951)
33. Wooldridge, M.: Reasoning about Rational Agents. The MIT Press, Cambridge (2000)