Chapter 1

DYNAMIC DESIRES

Mehdi Dastani, Zhisheng Huang, and Leendert van der Torre
Division of Mathematics and Computer Science
Vrije Universiteit Amsterdam
De boelelaan 1081a
1081 HV Amsterdam
The Netherlands
{mehdi,huang,torre}@cs.vu.nl

Abstract Logics of desires, preferences and goals have recently been proposed in planning and agent theory. In this paper we use a dynamic logic with utilitarian goals and desires to discuss the relation between goals, desires and utilities. From a static point of view goals are defined using desires and beliefs, and desires are defined using the utility function and a context. The consequences for the dynamics are that utilities are the most stable, then desires, and goals change most often. We illustrate the use of the logic to formalize certain aspects of negotiation. In particular, we show how one agent can influence the behavior of another agent by influencing his desires.

1. Introduction

Logics for desires and goals have been studied in several areas (see e.g. [RG91; Coh90; DT99; Kee76]). Both desires and goals play important roles in practical decision making. Here we understand desires as motivational forces inside agents, and goals as desirable states that the agents try to reach. They are understood as in the BDI process: desires are used to determine goals. A typical distinction is that an agent can choose his goals, whereas he cannot choose his desires. In this paper we formalize desires and goals in a decision-theoretic setting. An important issue in the logics of desires and goals is the relation between utility functions, desires, and goals. In decision theories as well as in BDI theories desires are usually formalized in terms of utility functions which implies that desires becomes as unchangeable as the utility function. This perspective is in line with the traditional perspective in economic and game theory, and has been questioned by e.g. Shoham [Sh97], where utility is interpreted as one which has close relationship with probabilities, and Castelfranchi [Cas98] who suggests that one can manipulate others' utilities. We are interested in formal models of *changing* desires and goals.

In this paper we look for a formal model in which the utility functions are typically constant, desires are relatively stable, whereas goals change much more frequently. For example, a person may want to go to a vacation island by airplane this week, but hears that the airplanes are in strike for the coming weeks. Since he can only go on holiday in the coming week, he changes his goal and now wants to stay at home and do house renovation. However, desires are much more stable. If an agent desires to go to the beach, then even if he cannot go because he has to work in the office, his desire does not disappear or change. Nevertheless, there are examples where desires change, and the typical example is the agent who stops to desire smoking once he falls in love with a girl that has lung cancer. It is not the case that he still likes to smoke but does not do it anymore; the cigarette no longer tastes. The change of the desire may be encoded by various factors. The issue has been discussed in the social sciences when researchers try to explain the effect of putting a nice looking girl on top of a car. This does not give us any new information about the car, but it nevertheless influences desires. The point is that agents can be influenced in many different ways, which do not operate only through their beliefs. This fundamental insight has already been part of the BDI model: goals are products of beliefs and desires.

Our representation is based on the simple model in figure 1.1. This figure should be read as follows. From decision theory we use the notion of a utility function. To define desire in terms of the utility function we have to solve the so-called lifting problem (to be explained later), for which we use a notion of *context*. To select desires which can function as goals we get into a second problem, the realizability problem: which goals can be realized? To solve this second problem we refer to the agent's beliefs. Since this second step has already been discussed extensively in the literature, for example in BDI logic, we focus in this paper on the first step.

Our aim is to develop a formal model which explains that the utility function is the most stable, then the desires and finally the goals. If the utility function changes then the other two change too, and if the desire changes then the goals change too. However, the inverse is not true. For example, goals can change when desires do not change because the

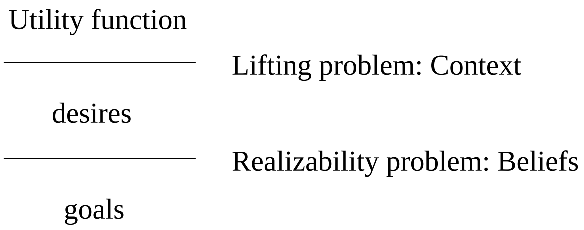

Figure 1.1. Utilities, desires and goals

agent's beliefs change. Moreover, the utility function may be constant whereas desires change, simply because the context changes.

This paper is organized as follows. In Section 2 we discuss the logic of utilities, desires and goals. In Section 3 we introduce the dynamic logic of utilitarian desires, and we show by an example with reasoning about desire in negotiation that utilities are more stable than desires. In Section 4, we discuss some properties of the logic and we discuss some future work. Finally, we conclude the paper in Section 5.

2. Utilities, desires and goals

> "The basic problem of social life among cognitive agents lies beyond mere coordination: *how to change the mind of the other agent? How to induce the other to believe and even to want something?* How to obtain that y does or does not do something? Of course, normally – but not necessarily – by communicating." [Cas98, p.170]

In this paper, desires of rational agents are assumed to reflect their utility function which in turn reflects their preferences. This relation formalizes the intuition that *rational agents desire to achieve states with the highest utility* when making decisions. This type of desire is called *utilitarian desire*; see [LvdTW; Per93; Bou94] for a further discussion on the relation between desires and decision making. Here utilitarian desire is understood as in classical decision theory: a rational agent acts as if he is maximizing expected utility based on an imaginary (or subjective) utility function and probability distribution. However, although we agree that desire and utility function are closely related concepts, we believe that they are distinct concepts since a utility function is assumed to remain constant while desires can change. Below, we try to elucidate this distinction by giving two examples. These examples illustrate how agents act to influence each other's desires and consequently affect the choice of actions to be taken by them.

2.1 Changing Desires: An Example

Consider again the agent who used to desire to smoke, but who no longer desires to do so since he has fallen in love with a girl that has lung cancer. If we consider the two propositions 'healthy' and 'smoking', then the following three states can be distinguished: 1) w_1: healthy and smoking, 2) w_2: healthy and not smoking, and 3) w_3: unhealthy and smoking. His utility function obeys the constraint $u(w_1) > u(w_2) > u(w_3)$. When the agent is not in love, then he focuses his attention to all states and he desires to smoke. But, when he falls in love with that girl, he focuses his attention to states w_2 and w_3 and he desires not to smoke.

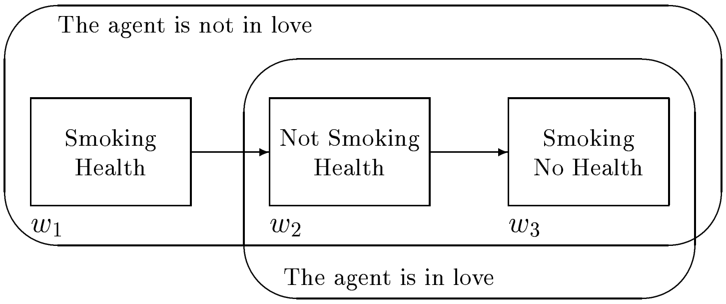

Figure 1.2. The smoking example.

Note that in this example, the utility function is defined on complete specifications (states) and desires on incomplete specifications (sets of states). To infer the desirability of a proposition the utility function defined on states has to be *lifted* to a desirability function on sets of states (called the lifting problem [Tho96]). The core of our idea is that although the utility function itself does not change, the lifting condition can change. Summarizing, the lifting condition in the first case (i.e. he is not in love) lifts the utility of the states to the desirability to smoke, and the lifting condition in the other case (i.e. he is in love) lifts the utility of the states to the desirability not to smoke.

2.2 The lifting problem

The main problem of defining utilitarian desires is that utility is defined on states, whereas desirability is defined on sets of states or propositions. This is called the *lifting problem*: the utility function on states has to be lifted to a desirability function on sets of states. For example, we can simply define the desirability of a set of states as the maximum of the utilities of its elements, or the lowest one, or some kind of average. As discussed in [vdTW00], the maximum reflects a kind of optimism that the best state will obtain, and the lowest a kind of pessimism that the worst state will obtain. The first is justified if we can determine or influence which of these states will obtain, and the latter is justified when an opponent can decide it (or when nature is seen as an opponent, like in Wald's criterion). In this paper we do not discuss the details of this particular choice and we therefore use the most popular choice, the maximum. Thus, a desire for p is satisfied by a utility function if the *relevant* state with the highest utility satisfies p, i.e. when for all *relevant* states w satisfying $\neg p$ there is a *relevant* state w' satisfying p such that the utility of w' is higher than the utility of w (the advantage of the latter definition is that it is more intuitive for infinite sets of states).

In this paper we are in particular interested in the question *which* states should be considered *relevant*. That is, we explicitly distinguish the following two stages in the lifting problem, of which in this paper we are primarily interested in the first one:

1 Select a relevant subset of all states. One constraint on the states is that the condition of the desire is true. Thus, for a desire "p if q" we select all q states. However, this is not all. For example, in the smoking example we restricted the focus first to states in which the agent is not in love, and thereafter to states in which the agent is in love.

2 Of these states, select the ones with the highest utility. A proposition p is desired if it is true in all these maximal utility states.

To solve the problem of state selection we introduce a rather abstract restriction of the set of states, inspired by recent work on modeling and using context [vdTTan00]. We therefore distinguish between two types of utilitarian desires,

ideal utilitarian desire, written $Id(p|q)$ for ideally p if q. It is defined in terms of utilities of all possible states, and

feasible utilitarian desire, written $D(p|q)$ for desirable p if q. It is defined in terms of utilities of all possible states in the agent's context.

2.3 Logic of desires and goals

In this section we formalize both types of utilitarian desires in Kripke models and in terms of possible states, their accessibilities, and their utilities. The starting point of our formalization is that the utility function is the same in all states. The following definition formalizes the pure utilitarian desire. An agent finds something ideal if and only if for all states in which it is false there is a state with a higher utility in which it is true. In this definition we use a mapping to the non-negative integers to ensure that there are no infinite descending chains.

Definition 1 *Let $\mathcal{M} = \langle W, U_i, C_i, N_i \rangle$ be a possible states model that consists of a set W of all possible states, a function U_i from W to the non-negative integers indicating the utility of states for agent i, an accessibility relation C_i that indicates which states the agent i considers relevant for his desires, and an accessibility relation N_i indicating the normality of states for agent i. A formula "ϕ is ideal if ψ" holds in state w for agent i (i.e. $\mathcal{M}, w \models Id_i(\phi|\psi)$) if and only if for all $\neg\phi \wedge \psi$ states there exists a $\phi \wedge \psi$ state with a higher utility, i.e.*

$$\begin{aligned}
\mathcal{M}, w \models Id_i(\phi|\psi) \quad &\iff \quad \forall w' \in W(\mathcal{M}, w' \models \neg\phi \wedge \psi) \exists w'' \in W \\
&\qquad (\mathcal{M}, w'' \models \phi \wedge \psi) \wedge (U_i(w'') > U_i(w'))
\end{aligned}$$

$$\begin{aligned}
\mathcal{M}, w \models D_i(\phi|\psi) \quad &\iff \quad \forall w' \in C_i(w)(\mathcal{M}, w' \models \neg\phi \wedge \psi) \exists w'' \in C_i(w) \\
&\qquad (\mathcal{M}, w'' \models \phi \wedge \psi) \wedge (U_i(w'') > U_i(w')) \\
\mathcal{M}, w \models G_i(\phi|\psi) \quad &\iff \quad \mathcal{M}, w \models D_i(\phi|\psi), \exists w' \in N(w) M, w' \models \phi \wedge \psi
\end{aligned}$$

Although desire and goal are defined as binary operator $D_i(\phi|\psi)$ and $G_i(\phi|\psi)$, the unary desire can be defined in terms of its binary counterpart as follows:

$$D_i\phi =_{def} D_i(\phi|\top) \qquad G_i\phi =_{def} G_i(\phi|\top)$$

If we allow conflicting desires (i.e. if we allow $D(p) \wedge D(\neg p)$ to be consistent), which is intuitive, [1] then we can derive $D(q)$ for any q. This is a consequence of the fact that $D(p)$ is a normal modal operator defined on normal possible state semantics. This means that the K-axiom and its derived conjunction rule (i.e. $\mathcal{M}, w \models (D_i p \wedge D_i q) \rightarrow D_i(p \wedge q)$) hold if $C_i(w)$ is always assumed to be a non-empty set. For simplicity, we do not allow conflicting desires and thus cannot satisfactorily reason with them.

Applying the conjunction rule to conflicting desires $D_i p \wedge D_i \neg p$ implies $D_i(p \wedge \neg p)$. However, it is possible to have conflicting desires if we do not require that $C_i(w)$ is always a non-empty set. Furthermore, note that the following weak version of conflicting desires is satisfiable:

$$D_i(\phi|\psi) \wedge D_i(\neg\phi|\rho).$$

The notion of context in the definition of feasible utilitarian desire is crucial. In fact, the context, which distinguishes the two types of desires, specifies which states are relevant (or feasible) for the agent's desires. It is this context that functions as a filter on possible states and determines the desirable states. This filtering function combined with the belief accessible states determine the goals accessible states. In general, given a set of formula specifying the set of possible accessible states, the context determines a set of desire formulae that specifies the desirable states. The set of desirable states is considered to be a subset of possible states. The set of desire formulae is then combined with the set of belief formulae, which result the set of formulae that specify the goal states.

Taking again the smoking example, it is now obvious that the change of desire from smoking to not smoking can be modeled by considering love as a determinating factor of the context: before love the context consisted of all three states and after falling in love the context contains only the state w_2 and w_3.

The reader may have noticed the close relation between our work and the BDI system. In order to illustrate this relation, consider the diagram illustrated in figure 1.3. A part of this diagram shows the general scheme of the BDI system consisting of beliefs, goals, and intentions related by dash arrows. According to this part of the diagram, desires determines goals from beliefs (this is indicated by the thick arrow) and commitments determines intentions from goals. Another part of this diagram

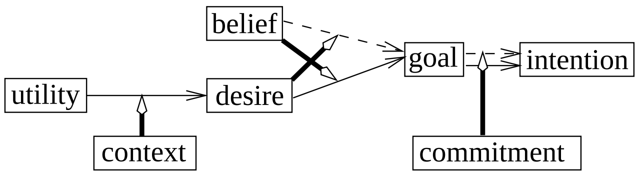

Figure 1.3. Enriching BDI with utilitarian desires.

shows a different path (characterized by solid arrows) starting from ideal utilitarian desire and ending at intentions. According to this part of the diagram, context determines (feasible utilitarian) desires from ideal utilitarian desires, beliefs determine goals from (feasible utilitarian) desires, and like in the BDI system, commitments determines intentions from goals.

3. An example from negotiation

In the remainder of this paper we are interested in modeling the agent's desires in negotiation. Rational agents in negotiation decide what action to take based on their desires which reflect their utility functions.

In this section we give an example to illustrate that the agent's desires are subject of change during negotiation. Moreover, since negotiation is usually modeled by game theory and because in game theory the utility function is assumed to remain constant during a game, we have to solve an apparent contradiction: on the one hand, it is reasonable to assume that the agent's desires can be changed during negotiation and on the other hand the utility function which is reflected by the rational agent's desire have to remain constant during negotiation. This apparent contradiction can be solved by *lifting* the utility function to a desirability function and allow the lifting condition to change on the basis of some context parameters.

Imagine a washing machine that wants to wash and thereby it must buy electricity from an electricity company. Figure 1.4 illustrates three possible washing states, w_1, w_2 and w_3 for the washing machine. In the state w_1, the washing machine washes in the afternoon and the electricity delivery is certain. Other states should be interpreted in similar way. Let the arrows between these three states denote the decrease of the utility of the states for the washing machine. Since w_1 provides the washing machine the highest utility, he desires to wash in the afternoon. However, the electricity provider desires to deliver electricity at night since they have usually more customers during the day. In this scenario, the electricity provider tries to influence the desires of the washing machine and informs him about the high prices of (certain deliverage) electricity for day consumption. The washing machine may not be able to spend much money for electricity. Consequently, the washing machine concentrates on states in which washing is cheap (i.e. w_2 and w_3) and because w_2 provides him more utility than w_3, he desires to wash at night. Note that the electricity provider may also change the focus of the washing machine in various ways such as making advertisements or refusing to sell electricity for the day consumption. In this paper, we do not consider what causes desires to change, and concentrate on modeling the dynamics of desires.

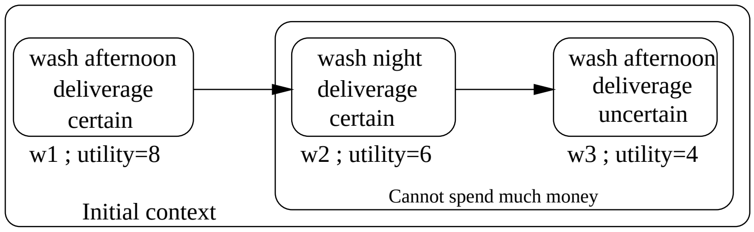

Figure 1.4. The washing machine example.

Note that the agent cannot decide to change his own desire to wash at night, like an agent can decide to change his adopted goal to wash at night. Desires are not consequences of the decision making process, they are their input. It is due to the influence from outside that desires are changed. Analogous examples can be found in pricing airline tickets, telephone calls, internet connection bandwidth, data storage, insurance contracts, and many other areas.

In order to model this type of negotiations, we represent the space of possible negotiation states and the accessibility relation defined on them as a Kripke frame. Note that the accessibility relation defined on negotiation states can be specified by negotiation actions such as propose, inform, and require actions. Moreover, we use dynamic logic to represent and reason about agent's negotiation strategies (actions). Since agents should reason about both their own desires as well as the desires of other agents we define agent's desire formally as a modal operator. This modal operator will be combined with dynamic operators resulting in a multi-modal logic. In this way, we can express the effects of actions on desires and enable agents to reason about desires and actions, and to decide on negotiation strategies.

We are now in position to show how agents can change their desires without changing their utility functions. This is because an agent's desires are now defined relative to the current agent's context and the fact that the current agent's context may change. Let UA, CN, and CA denote three washing states indicating washing in the afternoon with uncertain delivery of electricity, washing at night with certain delivery of electricity, and washing in the afternoon with certain delivery of electricity, respectively. Let also the utility of these states for the washing machine be: $UA = 4, CN = 6$, and $CA = 8$. Suppose that the initial context of the washing machine consists of all states. We assume that the washing machine can wash at one and only one period of the day in each state. In this initial context, the washing machine desires to wash in the afternoon since the only accessible state with highest utility is the one in which he washes in the afternoon. The washing machine receives now either the prices of the electricity company or become influenced by the advertisement announced by the electricity company. Consequently, the context of the washing machine may be changed such that it now consists of two states UA and CN. In this new context, the washing machine's desire has changed to wash at night since the only accessible state with the highest utility is the one in which he washes at night.

4. Dynamic logic

The machinery of dynamic logic, encompassing *choice* ($\cup$), *sequence* (;) and *iteration* (*), has been proposed in [BDvdT00] to reason in the context of negotiation protocols. Protocols are usually represented by finite state machines, which are equivalent in expressive power to the regular actions that are central to the ontology of dynamic logic. The *concurrency* operator $\cap$ is, for instance, used to model the reasoning of an agent that offers some service to different agents simultaneously. Action *negation* ($\sim$) is used in the formalization of the deontic notions 'obligation', 'permission' and 'prohibition', according to an approach first taken by Meyer [Mey85]. The *skip* denotes an action whose effect is void: doing a *skip* does not change the state we are in. It can be used in an expression like $T_{ij}(\alpha \leq skip)$ to express that by performing α our trust is decreased in comparison to the situation where we do nothing (we skip). The *any*-action refrains from action labels completely, which is useful if we want to reason about reachability of states, in a more temporal fashion. The *none*-action is the empty action. The most important elements of the logic introduced in [BDvdT00] are the modalities $C_{ij}(\alpha \leq \beta)$, whose intended meaning is 'agent i, is more *committed*, towards agent j, to perform α than to perform β', and $T_{ij}(\alpha \leq \beta)$ with the intended meaning 'the *trust* of agent i in agent j after performance of α is less than, or equal to the trust after performance of β'.

In this paper we extend dynamic logic with an operator D for desires and an operator B for beliefs, and we do not discuss commitments or trust. We (implicitly) assume that the utility function is global, i.e. the same in every state of the model. Thus, given utility function U_i for agent i, the fact that the utility function is global can be formally expressed by $U_i(\phi) \to [\alpha]U_i(\phi)$ which states that the utility function is not changed by any action. Since we are not interested in the desire operator by itself, but only in the interaction between actions and desires, we use a simple normal modal logic to formalize them. The belief operator is introduced in the logic, so that we define the goal operator G in terms of desires and beliefs, by which we can discuss the relations between desires and goals.

Given a set $\mathcal{A}$ of action symbols and $a \in \mathcal{A}$, a set $\mathcal{P}$ of proposition symbols and $p \in \mathcal{P}$, a ***well formed formula*** ϕ is defined through the following BNF:

$$\begin{array}{lll} \phi & ::= & p \mid \neg\phi \mid \phi \vee \psi \mid \langle\alpha\rangle\phi \mid I_i(\phi|\psi) \mid D_i(\phi|\psi) \mid B_i\phi \\ \alpha & ::= & a \mid any \mid none \mid skip \mid \ \sim\alpha \mid \alpha \cup \alpha' \mid \alpha \cap \alpha' \mid \alpha;\alpha' \mid \alpha^* \end{array}$$

The semantics is defined in a fairly standard way.

Definition 2 *Given a set $\mathcal{A}$ of action symbols, a set $\mathcal{P}$ of proposition symbols, a structure $\mathcal{S} = (S, R_{\mathcal{A}}, \pi, U_i, C_i, B_i)$ is defined as:*

- *S is a nonempty set of possible states*
- *$R_{\mathcal{A}}$ is an $\mathcal{A}$-indexed collection of (reachability) relations over $S \times S$*
- *π is a valuation function $\pi : \mathcal{P} \to 2^S$ that interprets propositions $p \in \mathcal{P}$*
- *U_i is a total ordering over S that indicates the utility of states for agent i*
- *C_i is an accessibility relation that indicates which states the agent i considers relevant for his desires.*
- *B_i is a serial accessibility relation that indicates which states the agent i considers possible.*

Definition 3 *The meaning of well-formed formulas in a state s of a structure $\mathcal{S}$ is given by:*

$$
\begin{array}{lcl}
R_\alpha & = & R_a \text{ for } \alpha = a \text{ for } a \in \mathcal{A} \\
R_{none} & = & \emptyset \\
R_{skip} & = & \{\langle s, s\rangle : \text{ for all } s \in S\} \\
R_{\alpha \cap \alpha'} & = & R_\alpha \cap R_{\alpha'} \\
R_{\alpha \cup \alpha'} & = & R_\alpha \cup R_{\alpha'} \\
R_{any} & = & (R_{a \cup b \cup c \ldots})^* \text{ with } \{a, b, c, \ldots\} = \mathcal{A} \\
R_{\sim\alpha} & = & R_{any} \setminus R_\alpha \\
R_{\alpha;\alpha'} & = & R_\alpha \circ R_{\alpha'} \\
R_{\alpha^*} & = & (R_\alpha)^*
\end{array}
$$

$$
\begin{array}{lll}
\mathcal{S}, s \models P & \text{iff} & s \in \pi(P) \\
\mathcal{S}, s \models \neg\phi & \text{iff} & \text{not } \mathcal{S}, s \models \phi \\
\mathcal{S}, s \models \phi \wedge \psi & \text{iff} & \mathcal{S}, s \models \phi \wedge \mathcal{S}, s \models \psi \\
\mathcal{S}, s \models \langle\alpha\rangle\phi & \text{iff} & \exists s' \in S, \ s.t. \ \langle s, s'\rangle \in R_\alpha \wedge \mathcal{S}, s' \models \phi \\
\mathcal{S}, s \models I_i(\phi|\psi) & \text{iff} & \forall s' \in S(\mathcal{S}, s' \models \neg\phi \wedge \psi) \\
 & & \exists s'' \in S(\mathcal{S}, s'' \models \phi \wedge \psi) \wedge (U_i(s'') > U_i(s')) \\
\mathcal{S}, s \models D_i(\phi|\psi) & \text{iff} & \forall s' \in C_i(s)(\mathcal{S}, s' \models \neg\phi \wedge \psi) \\
 & & \exists s'' \in C_i(s)(\mathcal{S}, s'' \models \phi \wedge \psi) \wedge (U_i(s'') > U_i(s')) \\
\mathcal{S}, s \models B_i\phi & \text{iff} & \forall s' \in S, s.t. \langle s, s', \rangle \in B_i \wedge S, s' \models \phi
\end{array}
$$

For the discussion in this paper, in which we are interested in the interaction between the two types of modal operators, the exact definition of the operators is not relevant. We define $[\alpha]\phi$ as $\neg\langle\alpha\rangle\neg\phi$. Moreover, the goal operator G is defined in terms of desires and beliefs as follows. ϕ is a goal for agent i (under the condition ψ) if and only if ϕ is a desire for i and the agent i does not believe that ϕ is impossible (under the condition ψ).

$$G_i(\phi|\psi) \Leftrightarrow D_i(\phi|\psi) \wedge \neg B_i \neg(\phi \wedge \psi).$$

The unary operator of the goal is defined as:

$$G_i\phi \Leftrightarrow G_i(\phi|\top).$$

Thus, we have the following relation between goals and desires:

$$G_i\phi \leftrightarrow D_i\phi \wedge \neg B_i \neg \phi.$$

4.1 The example

The following example illustrates how desires change in a dynamic situation such as the negotiation setting. Consider the following Kripke model $\mathcal{S}$, illustrated in Figure 1.5, representing one possible step in the negotiation between the washing machine and the electricity provider. In

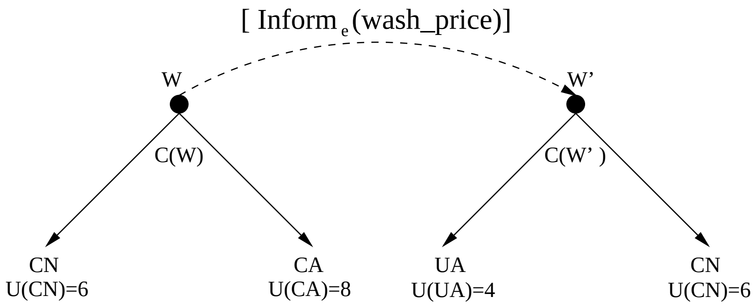

Figure 1.5. A model $\mathcal{S}$ for the washing machine.

this model there are two types of relations defined on states. One relation is represented by dashed lines and they indicate possible negotiation actions, i.e. the relation specifies the space of possible negotiation actions (extensive negotiation game). The second relation is represented by solid lines and they indicate accessible states of an agent, i.e. the relation expresses the context of an agent.

In this model, the dash line indicates a proposal initiated by the electricity provider, i.e. $Propose_e(Wash(night))$, and the solid lines from each state indicates the context of an agent in that state. It is this context which determines the desires of agents. For instance, in the state w of model $\mathcal{S}$, the washing machine desires to wash in the afternoon since the only accessible state with the highest utility in the context of w for the washing machine is CA, i.e. $\mathcal{S}, w \models D_{wm}(Wash(afternoon)|\top)$. Suppose, after the proposal done by the electricity provider the context for the washing machine changes such that he desires now to wash at night, i.e. $\mathcal{S}, w' \models D_{wm}(Wash(night)|\top)$. Note that although the utility of the washing machine for states are not changed, his desire to wash in the afternoon is changed.

4.2 Theorems

The following formulas are theorems of the logic.

Stability

$Id_i(\phi|\psi) \rightarrow [\alpha]Id_i(\phi|\psi)$.
(Actions do not change pure utilitarian desires at all.)

Left Expansion

$Id_i(\phi|\psi) \rightarrow Id_i(\phi \wedge \psi|\psi)$.
$D_i(\phi|\psi) \rightarrow D_i(\phi \wedge \psi|\psi)$.
$G_i(\phi|\psi) \rightarrow G_i(\phi \wedge \psi|\psi)$.

Right Expansion

$Id_i(\phi|\psi) \rightarrow Id_i(\phi|\phi \wedge \psi)$.
$D_i(\phi|\psi) \rightarrow D_i(\phi|\phi \wedge \psi)$.
$G_i(\phi|\psi) \rightarrow G_i(\phi|\phi \wedge \psi)$.

Consistency

$\neg Id_i(\phi|\neg\phi)$ if $\neg\phi$ satisfiable.
$\neg D_i(\phi|\neg\phi)$ if $\neg\phi$ satisfiable.
$\neg G_i(\phi|\neg\phi)$ if $\neg\phi$ satisfiable.
$\neg B_i\bot$.
$\neg G_i\bot$.

Conjunction Decomposition

$Id_i(\phi \wedge \rho|\psi) \leftrightarrow Id_i(\phi|\psi) \wedge Id_i(\rho|\psi)$.
$D_i(\phi \wedge \rho|\psi) \leftrightarrow D_i(\phi|\psi) \wedge D_i(\rho|\psi)$.
$G_i(\phi \wedge \rho|\psi) \leftrightarrow G_i(\phi|\psi) \wedge G_i(\rho|\psi)$.
$B_i(\phi \wedge \rho) \leftrightarrow B_i(\phi) \wedge B_i(\rho)$.

Dynamics

$G_i(\phi|\psi) \wedge [a]B_i\neg(\phi \wedge \psi) \rightarrow \neg\langle a\rangle G_i(\phi|\psi)$.
$\langle a\rangle B_i(\phi \wedge \psi) \wedge [a]\neg G_i(\phi|\psi) \rightarrow \langle a\rangle\neg D_i(\phi|\psi)$.

4.3 Classification of actions

The formal language combines dynamic action operators with the desire operator. This combined language is useful to characterize several properties. First we distinguish between several types of formulas:

objective formulas do not contain any modality

static formulas do not contain a dynamic modality

desire atom starts with the operator D or the operator I

desire formula is a combination of conjunction, disjunction and desire atoms

Some formulas hold for certain types of actions, and thus *characterize* these actions. For example:

pure actions α do not change pure utilitarian desires at all and are characterized by $I_i(\phi|\psi) \rightarrow [\alpha]I_i(\phi|\psi)$

influence actions α only change desires and are characterized by $\phi \rightarrow [\alpha]\phi$ if ϕ is objective

fulfillment actions α only change facts if this is desirable and are characterized by $\phi \rightarrow ([\alpha_i](\neg\phi \wedge \psi) \rightarrow D_i(\neg\phi|\psi))$

5. Future Research

In this paper we focused on motivation and semantics. However, if agents reason about negotiation they may prefer to use the logical proof theory. For example, the agent can reason about the other agent's desires and decide how to influence them and which negotiation action to take. Thus there is a need for an axiomatization.

Besides, the logical language may be useful to distinguish different concepts. For example, we can classify actions in actions that only change facts, actions that only change desires, and actions that change both. Moreover, actions can be partitioned in actions that fulfill desires and actions that do not. This can be the basis of agent profiles: an agent that often does what you desire may become one of your friends.

6. Concluding remarks

In this paper we studied utilitarian desires in a dynamic environment. We proposed a semantics for a dynamic logic of context dependent desires. The context formalizes the intuition that desires can change whereas the utility function remains constant. We have shown that the notion of context involved is different from notions of belief or knowledge which have been used in a decision-theoretic setting. Technically, the work is related to dynamic epistemic logic. In these studies, a distinction is made between general actions and observation actions. In the former the agents state changes, whereas in the latter only the agent's beliefs change.

This work is presented in the context of practical negotiations where agents can reach agreements by influencing other agent's desires. In the negotiation context, the interaction between dynamic operators and the desire operator is illustrated by the effect of negotiation actions (dynamic operators) on an agent's desire (desire operator).

Notes

1. Intuitively, one can for example desire to be rich and not-working although not-working implies not-rich.

References

J. Bell and Z. Huang. *Dynamic goal hierarchies.* In: L. Cavedon, A. Rao, W. Wobcke (eds.), Intelligent Agent Systems, Theoretical and Practical Issues, LNAI 1209, Springer, 1997, 88-103.

A. Rao and W. Wobcke (eds.). *Intelligent Agent Systems, Theoretical and Practical Issues.* LNAI 1209, Springer, 1997, 88-103.

J. Broersen, M. Dastani, and L. van der Torre. *Levelled commitment and trust in negotiation.* In Proceedings of the Autonomous Agents 2000 Workshop on Deception, Fraud and Trust in Agent Societies, Barcelona, 2000.

K. Binmore. *Fun and Games: A text on Game theory.* D.C. Heath and Company Lexington, Massachusetts Toronto 1992.

C. Boutilier. *Toward a logic for Qualitative Decision Theory.* Proceedings of the 4th International Conference on Principles of Knowledge Representation and Reasoning (KR'94),75–86, 1994.

M. Bratman. *Intentions, Plans, and Practical Reason.* Harvard University Press, Cambridge, MA, 1987.

C. Castelfranchi. *Modelling Social Action for AI Agents.* Artifical Intelligence 103(1-2): 157-182, 1998.

P. Cohen and H. Levesque. *Intention is choice with commitment.* Artificial Intelligence, 42:213-261, 1990.

J. Doyle and R. Thomason. *Background to qualitative decision theory.* In: AI magazine 20(2):55-68, summer 1999.

R.L. Keeney and H. Raiffa. *Decisions with Multiple Objectives: Preferences and Value Trade-offs.* Wiley and Sons, New York, 1976.

J. Lang, L. van der Torre and E. Weydert. *Utilitarian Desires.* Autonomous Agents and Multi Agent Systems, To appear.

J.-J.Ch. Meyer. *A different approach to deontic logic: deontic logic viewed as a variant of dynamic logic.* Notre Dame Journal of Formal Logic, 29:109–136, 1985.

M. Osborne and A. Rubinstein. *A course in Game Theory.* MIT press, 1994.

J. Pearl. *From conditional oughts to qualitative decision theory.* Proceedings of the Ninth Conference on Uncertainty in Artificial Intelligence (UAI'93), 12–20, 1993.

H. Raiffa. *The art and science of negotiation.* Harvard university press, Cambridge, Mass., 1982.

A. Rao and M. Georgeff. *Modeling rational agents within a bdi architecture.* Proceedings of the 1st International Conference on Principles of Knowledge Representation and Reasoning (KR'91), 1991.

A. Rao and M. Georgeff. *An abstract architecture for rational agents.* Proceedings of the 2th International Conference on Principles of Knowledge Representation and Reasoning (KR'92), 1992.

Y. Shoham. *Conditional Utility, Utility Independence, and Utility Networks.* Proceedings of the Thirteenth Conference on Uncertainty in Artificial Intelligence (UAI'97), 429-436, 1997.

R. Thomason and R. Horty *Nondeterministic action and Dominance: foundations for planning and qualitative decision.* Proceedings of the Sixth Conference on Theoretical Aspects of Rationality and Knowledge. De Zeeuwse Stromen, The Netherlands, March 17-20 1996. Morgan Kaufmann Publishers, 229–250, 1996.

L. van der Torre and Y. Tan. *Contextual Deontic Logic: violation contexts and factual defeasibility.* P. Bonzon, M. Cavalcanti and R. Nossum, eds, Formal Aspects of Context, Applied Logic Series, 143-166, Kluwer, 2000.

L. van der Torre and E. Weydert. em Parameters for utilitarian desires in a qualitative decision theory. Applied Intelligence, 14: 285-301, 2001.