Dynamic Desires

Mehdi Dastani, Zhisheng Huang, and Leendert van der Torre

Department of Artificial Intelligence, Vrije Universiteit
De Boelelaan 1081a, 1081 HV Amsterdam, the Netherlands
{mehdi,huang,torre}@cs.vu.nl
http://www.cs.vu.nl/~{mehdi,huang,torre}

Abstract. Logics of desires, preferences and goals have recently been proposed in planning and agent theory. In this paper we introduce a dynamic logic with utilitarian desires and we discuss the relation between desires and utilities. Given that an agent's desires refer to his utility function, which we assume to be constant, we resolve the paradox that the stability of the agent's utility function does not imply the stability of his desires. 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

Autonomous agents reason frequently about preferences [Kee76] such as desires and goals. Recently several logics of desires have been proposed in two areas:

Planning. Work in qualitative decision theory [DT99] illustrates how planning agents are provided with goals – determined by rational desires – and charged with the task of discovering (or performing) some sequence of actions to achieve those goals. In qualitative decision theory desires are usually formalized as constraints on qualitative abstractions of utility functions.

Agent theory. Work on cognitive and social agents [Cas98] illustrates how autonomous agents reason with and about cognitive concepts. For example, Cohen and Levesque [Coh90] explore principles governing the rational balance among an agent's beliefs, goals, actions and intentions, Rao and Georgeff [RG91, RG91, RG92] show how different rational agents can be modeled by imposing certain conditions on the persistence of an agent's beliefs, goals and intentions (the BDI model). In some agent theories, desire is formalized as a modal operator.

In this paper we study the combination of desires with action operators. Multi-modal logics introduce complexities not found in the modal approaches, due to interaction of the modalities. An important issue in the logics of preferences, desires and goals is the relation between utility functions, desires, and goals. In our approach, we assume that the utility function is constant, desires can change but are relatively stable, and goals can change together with the agent's commitments.

Utility functions and desires. The perspective on utility functions is in line with the traditional perspective in economic theory. It has been questioned by e.g. Shoham [Sh97], where utility is interpreted as one which has close relationship with probabilities. In [Cas98], Castelfranchi suggests that one can manipulate others' utilities, but in this paper we argue that the counter-arguments to a fixed utility function originate in a confusion between utility function and desires.

Desires and goals. Agent's goal are determined by agent's desire, i.e. goals are rational (chosen) desires [BH97]. Goals are related to the agent's planning, and therefore desires are useful for agent's planning.

Given the combination of desires with action operators in this paper, we have to explain the following apparent contradiction: on the one hand, desires can be influenced by other agents [Cas98], and on the other hand desires reflect utility functions which are not subject of change. The answer given in this paper is the following: a utility function is defined on complete specifications (states) and desires on incomplete specifications (sets of states or propositions). 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). Though the utility function itself cannot change, the lifting condition can change. This solution of the paradox is illustrated in the examples of the following section and discussed in more detail later in this paper.

This paper is organized as follows. In Section 2 we discuss a motivating example with reasoning about desires in negotiation. In Section 3 we present the logic, which is illustrated by an example in Section 4.

2 Two Motivating examples

> 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]

We consider utilitarian desires: an agent desires outcomes with high utility [LvdTW]. 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. In this sense, agent's decision-making depends on his desires. In order to predict other agent's actions an agent may consider the other agent's desires instead of their utility function. Yet, in more complicated scenarios agents may act to influence each other's desires and consequently affect the choice of actions to be taken by them. Analogously, agents may inform each other about opportunities and try to persuade each other to change or modify their initial *adopted goals*. However, in this paper we restrict ourselves to desires.

Example 1 (Washing Machine).

Imagine a washing machine that wants to wash and thereby he must first buy electricity from an electricity company. Figure 1 illustrates three possible washing states, w_1, w_2 and w_3 for the washing machine. In state w_1, the washing machine washes in the afternoon, washing is expensive, and the electricity deliverage is certain. Other states should be interpreted in similar way. Let the arrows between these three states denote the decrease of 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 the high prices of electricity for day consumption. The washing machine may reason and conclude that he cannot spend much money and that he should better use cheaper electricity. Consequently, the washing machine will concentrate 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 by making advertisements.

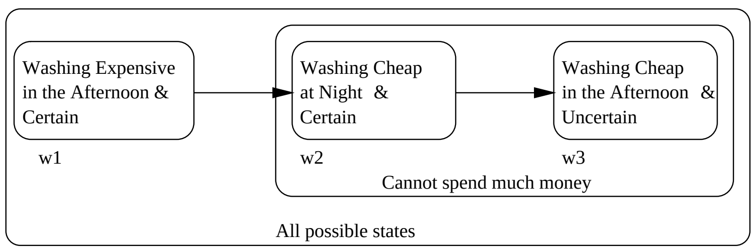

Fig. 1. The washing machine example.

Note that the agent cannot decide to change his own desire to wash during the night, like an agent can decide change his adopted goal to wash during the 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 area's.

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 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.

Example 2 (Smoking).
Consider an agent who used to desire to smoke, but who no longer desires to do so since he heard that smoking endangers his health. His desire to smoke thus changed after he learned the health implications of smoking. We distinguish the following four cases:

1. smoking is not bad for his health and he smokes.
2. smoking is not bad for his health and he does not smoke.
3. smoking is bad for his health and he does not smoke.
4. smoking is bad for his health and he smokes.

However, we do not use a proposition 'smoking is bad for agent's health' but a simpler one denoting whether the agent is healthy. For simplicity, we assume that if smoking is bad for the health, then smoking *causes* unhealthy. If we consider the two propositions 'healthy' and 'smoking', then the above four cases can be represented by the following four states. His utility function obeys the constraint $u(w_1) > u(w_2) > u(w_3) > u(w_4)$:

1. w_1: healthy and smoking (case 1.)
2. w_2: healthy and not smoking (case 2. and 3.)
3. w_3: unhealthy and smoking (case 4.)
4. w_4: unhealthy and not smoking (not considered)

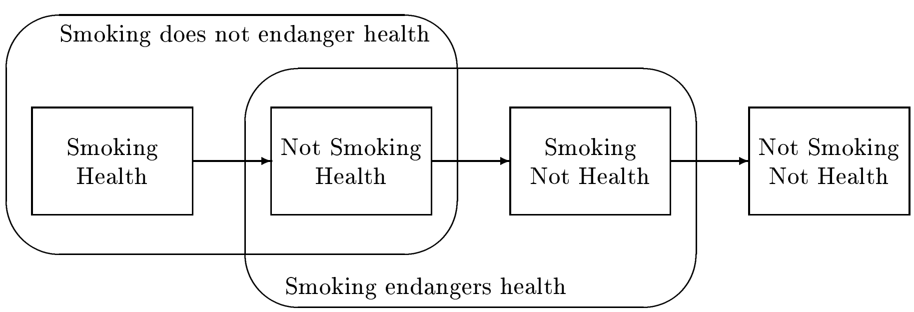

Fig. 2. The smoking example.

When the agent believes that smoking is not bad for his health, then he focuses his attention to the first two states and he desires to smoke. Otherwise, he focuses his attention on the middle two states and he desires not to smoke. Summarizing, the lifting condition in the first case (i.e. the believe that smoking in not bad for health) lifts the utility of the first states to the desirability to smoke, and the lifting condition in the other case (i.e. the believe that smoking in bad for health) lifts the utility of the third state to the desirability to smoke.

Note that in this case the desirability to smoke coincides with the *expected* utility to smoke, because in both cases only one state is assumed to be possible. That is, the conditional expected utility of a proposition s, written as $EU(s)$, is

$$EU(s) = \frac{\Sigma_{w \models s}(u(w) \times p(w))}{\Sigma_{w \models s}(p(w))}$$

When smoking does not endanger health then $EU(smoking) = u(w_1)$ because only w_1 and w_3 are smoking states and moreover $p(w_3) = 0$. Likewise, when smoking does endanger health then $EU(smoking) = u(w_3)$ because only w_1 and w_3 are smoking states and moreover $p(w_1) = 0$. Desirability and expected utility are both defined on propositions or sets of states, but they are something completely different. The former expresses a wish that some state occurs, whereas the latter expresses the desirability to do some *action* with expected outcomes reflected by the probability distribution used in calculating the expected utility.

3 Utilitarian desires

In this paper, desires of rational agents are assumed to reflect his utility function which in turn reflects his 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] for a further discussion on the relation between desires and decision making.

In particular, we are interested in modeling agent's desire in negotiation. Rational agents in negotiation decide what action to take based on their desires which reflect their utility function. In previous sections, we have argued that agent's desire are subject of change during negotiation and we have giving some examples to illustrate this claim. 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 agent's desires can be changed during negotiation and on the other hand the utility function which is reflected by 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.

3.1 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 smoking is healthy, and thereafter to states in which smoking is unhealthy. Moreover, in the washing machine example the focus on states is changed when the belief of the washing machine regarding his financial ability is changed (note that another reason to change the focus on states was the advertisement effects).
2. Of these states, select the ones with highest utility. A proposition p is desired if it is true in all these maximum utility states.

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

pure utilitarian desire , written $I(p|q)$ 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 agents context.

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.

Definition 1. *Let $\mathcal{M} = \langle W, U_i, C_i \rangle$ 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 reals indicating the utility of states for agent i, and an accessibility relation C_i that indicates which states the agent i considers relevant for his desires. A formula "ϕ is ideal if ψ" holds in state w for agent i (i.e. $\mathcal{M}, w \models I_i(\phi|\psi)$) if and only if for all $\neg\phi \wedge \psi$ states there exists a $\phi \wedge \psi$ state with a higher utility, i.e.*

$$\mathcal{M}, w \models I_i(\phi|\psi) \iff \forall w' \in W (\mathcal{M}, w' \models \neg\phi \wedge \psi) \exists w'' \in W \\ (\mathcal{M}, w'' \models \phi \wedge \psi) \wedge (U_i(w'') > U_i(w'))$$

$$\mathcal{M}, w \models D_i(\phi|\psi) \iff \forall w' \in C_i(w) (\mathcal{M}, w' \models \neg\phi \wedge \psi) \exists w'' \in C_i(w) \\ (\mathcal{M}, w'' \models \phi \wedge \psi) \wedge (U_i(w'') > U_i(w'))$$

Note that 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 agent's desires. We follow the idea of epistemic logic that distinct states are technical implementation of distinct knowledge and reformulate the question "which states are relevant" as "*which* kinds of knowledge is considered". In other words, we reformulate the question as which kinds of knowledge specify the context. In the following definition background knowledge, factual knowledge and feasibility knowledge are distinguished. Note that in many BDI logics the epistemic belief operator B is used for all three types of knowledge [1].

Definition 2 ([LvdTW]).

- *Background knowledge is knowledge which is true in all states considered possible,*
- *factual knowledge is knowledge about which state is the actual state, and*
- *feasibility knowledge is knowledge about which state can be reached from the actual state.*

Example 3. In the smoking example, it seems intuitive to say that the belief that smoking is unhealthy is a kind of background knowledge. In the washing machine example his (new) belief that he cannot pay the required electricity prices is assumed to be more like factual knowledge. In the same example we also argued that the desire of the washing machine may have been changed because of the advantages to wash at night advertised by the electricity provider. In fact the advertisement induces the knowledge that some states are attractive. This type of knowledge may be considered as feasibility knowledge.

[1] However, in this paper, we do not introduce the belief operator in the model. We would not investigate the belief revision as well. In this paper, we just consider the cases in which belief revision result from agents' actions, in particular, from the negotiation actions. So, using dynamic operators would be powerful enough for the formalization. Note that the introduction of a belief operator *by itself* does not solve our problems, because it may easily be that an agent desires something which he does not believe (i.e. he may desire to go to the dentist without feeling pain though he believes that going to the dentist implies pain).

We think, however, that these decision-theoretic types of knowledge are not sufficient to explain the wide range of examples about desires. We may desire things which are not considered possible. I may desire to go to the dentist without feeling pain, although I know that going to the dentist implies pain [Bra], and I may desire to live without dying, though I know that one day I will die. Despite this knowledge, I can still *imagine* a state in which going to the dentist does not imply pain, or a state in which we do not die.

We are now in position to show how agents can change desires without changing utility function. This is because agent's desires are now defined relative to the current agent's context and the fact that the current agent's context may change.

Example 4. Let EN, CN, EA, and CA be four washing states indicating washing expensive at night, washing cheap at night, washing expensive in the afternoon, and washing cheap in the afternoon, respectively. Let also the utility of these states for the washing machine be as follows: $EN = 2, CN = 4, EA = 6$, and $CA = 8$. Suppose that the current context of the washing machine consists of two states CN and EA. We assume that the washing machine can wash at one and only one period of the day in each state, i.e. washing at night in one state implies not washing in the afternoon in that state. In this current 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 EN and CN. In this new context, the washing machine will change its desire to wash at night since the only accessible state with highest utility is the one in which he washes at night

The example can be futher analyzed by introducing concepts from multi attribute utility theory [Kee76], in which different attributes are distinguished. In the washing machine example there are two atributes: the price and the time of washing. Objectives are to wash cheaply and to wash during the day. Typically objectives conflict, like in our example. The independence between different attributes can be represented by independence statements, which can lower the complexity of the utility function. The independence statements correspond to a notion of ceteris paribus preferences. In this paper we do not further discuss the extension with attributes, independence statements, ceteris paribus preferences etc because it is irrelevant for the dynamics of desires, and therefore it is irrelevant for our problems.

3.2 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 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 $U_i(\phi) \rightarrow [\alpha]U_i(\phi)$ 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.

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:

$$\phi ::= p \mid \neg\phi \mid \phi \vee \psi \mid \langle\alpha\rangle\phi \mid I_i(\phi|\psi) \mid D_i(\phi|\psi)$$
$$\alpha ::= a \mid any \mid none \mid skip \mid \sim\alpha \mid \alpha \cup \alpha' \mid \alpha \cap \alpha' \mid \alpha ; \alpha' \mid \alpha^*$$

The semantics is defined in a fairly standard way.

Definition 3. *Given a set $\mathcal{A}$ of action symbols, a set $\mathcal{P}$ of proposition symbols, a structure $\mathcal{S} = (S, R_{\mathcal{A}}, \pi, U_i, C_i)$ 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 $\pi : \mathcal{P} \rightarrow 2^S$ 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.*

Definition 4. *The meaning of well-formed formulas in a state s of a structure $\mathcal{S}$ is given by:*

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

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

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 $\neg\langle\alpha\rangle\neg\phi$.

Example 5. Consider the following Kripke model $\mathcal{S}$, illustrated in Figure 3, representing one possible step in the negotiation between the washing machine and the electricity provider. In this model there are two types of relations defined

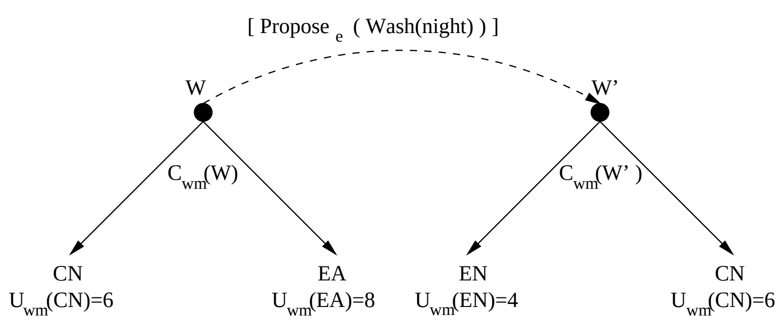

Fig. 3. A model $\mathcal{S}$ for the washing machine.

on states. One relation is represented by dash 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 expensive in the afternoon since the only accessible state with the highest utility in the context of w for the washing machine is EA, i.e. $\mathcal{S}, w \models D_{wm}(Wash(afternoon)|expensive(afternoon))$. After the proposal done by the electricity provider the context for the washing machine will be changed such that he desires to wash cheap at night, i.e. $\mathcal{S}, w' \models D_{wm}(Wash(night)|cheap(night))$. Note that although the utility of the washing machine for states are not changed, his desires are changed.

4 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.

In this section we do not discuss these issues in any depth, but we give a few examples of the kind of proof theoretic extensions we have in mind.

4.1 Axioms

Some axioms can be accepted for all types of actions and desires, for example the exchange between dynamic and desire operators.

Stability

$I_i(\phi|\psi) \rightarrow [\alpha]I_i(\phi|\psi)$

(Actions do not change pure utilitarian desires at all.)

Left Expansion

$I_i(\phi|\psi) \rightarrow I_i(\phi \wedge \psi|\psi)$

$D_i(\phi|\psi) \rightarrow D_i(\phi \wedge \psi|\psi)$

Right Expansion

$I_i(\phi|\psi) \rightarrow I_i(\phi|\phi \wedge \psi)$

$D_i(\phi|\psi) \rightarrow D_i(\phi|\phi \wedge \psi)$

Consistency

$\neg I_i(\phi|\neg\phi)$.

$\neg D_i(\phi|\neg\phi)$.

Conjunction Decomposition

$I_i(\phi \wedge \rho|\psi) \rightarrow I_i(\phi|\psi) \wedge I_i(\rho|\psi)$

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

4.2 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 and 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
fulfilment actions 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 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 agents 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 agents desire (desire operator).

References

[BH97] Bell, J., and Huang, Z., Dynamic goal hierarchies, in: L. Cavedon, A. Rao, W. Wobcke (eds.), *Intelligent Agent Systems, Theoretical and Practical Issues* , LNAI 1209, Springer, 1997, 88-103.

[BDvdT00] J. Broersen, M. Dastani, L. van der Torre. Levelled commitment and trust in negotiation. In: *Fraud2000 Workshop*. 2000.

[Bin92] K. Binmore, editor. Fun and Games. *A text on Game theory*, Heath, 1992.

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

[Cas98] C.Castelfranchi. Modelling Social Action for AI Agents, *Artifical Intelligence* 103(1-2): 157-182, 1998.

[Coh90] Cohen and Levesque. Intention is choice with commitment, *Artificial Intelligence*:42, 1990.

[DT99] Doyle and Thomason. Background to qualitative decision theory. In: *AI magazine* (20:2) summer 99. pages 55-68.

[Kee76] R.L. Keeney and H. Raiffa. *Decisions with Multiple Objectives: Preferences and Value Trade-offs.* Wiley and Sons, New York, 1976.

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

[Mey85] 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.

[OR94] M. Osborne and A. Rubinstein. A course in Game Theory. *MIT press*, 1994.

[Rai82] H. Raiffa. *The art and science of negotiation.* Harvard university press, Cambridge, Mass., 1982.

[RG91] Rao and Georgeff. Modeling rational agents within a BDI architecture, In: *Proceedings of the KR91*

[RG91] Rao and Georgeff. Modeling rational agents within a bdi architecture. In *Proceedings of the KR91*, 1991.

[RG92] Rao and Georgeff. An abstract architecture for rational agents. In *Proceedings of the KR92*, 1992.

[Sh97] Shoham, Y., Conditional Utility, Utility Independence, and Utility Networks. Proceedings of UAI97, 429-436, 1997.

[vdTW00] L. van der Torre and E. Weydert. Parameters for utilitarian desires in a qualitative decision theory. *Applied Intelligence*. To appear.