Tests as Epistemic Updates
Pursuit of Knowledge*

B. van Linder   W. van der Hoek   J.-J. Ch. Meyer[†]
Utrecht University
Department of Computer Science
P.O. Box 80.089
3508 TB Utrecht
The Netherlands
Email: bernd@cs.ruu.nl

Abstract

In this paper we look into the sort of tests that, in contrast with the tests in dynamic logic, can be used to acquire knowledge. Intuitively, these tests can be thought of as observations that provide information to the person that makes the observation. We give three demands that actions should meet for them to rightly be called tests: informativeness, truthfulness, and minimal change. The first of these demands formalizes the intuition that tests should provide information concerning the proposition that is tested, the second states that the truth or falsity of the proposition that is tested may not be affected, and the third postulate states that tests should not cause more changes to the world than necessary. We show how a general test action can be incorporated in a framework that formalizes (part of) the behaviour of rational agents. To this end notions from dynamic and epistemic logic are combined with some aspects of update semantics. We give two implementations of tests, one in a modal **KT** setting and one in an **S5** setting. Both implementations are proved to satisfy the Gärdenfors postulates for belief expansion.

Content Areas: Belief Revision, Common Sense Reasoning, Reasoning about Action.

1. Introduction

Tests, and the act of testing, play a natural part in everyday human life. People that want to get their driver's license take driving tests; during medical check ups hearing tests and eye tests are taken; applicants for a vacancy take psychological and/or i.q. tests; would-be athletes take 12-minute running tests and famous athletes are dope tested.

Also in the field of AI that deals with the formalization of the behaviour of rational agents[1] a natural part is played by testing. By testing, robots can acquire knowledge that might be necessary to guide their actions and that is used in their planning.

In this paper we consider a test to be some kind of unspecified procedure that can be used to acquire knowledge concerning the truth or falsity of certain propositions. For instance, an eye test yields knowledge concerning the truth or falsity of the proposition 'is the patient's eyesight satisfactory', and a driving test does the same thing for the proposition 'is the candidate fit to drive a car'. We will not consider *how* the actual testing is performed, i.e, in case of an eye test, it is of no concern to us how exactly the quality of the patient's eyesight is determined as long as knowledge is acquired. As such, these tests can naturally be seen as observations: an eye tests is then considered to be an observation concerning the sight of the patient, and by a psychological test the capacities of some applicant are observed.

1.1. Organization of the paper

The rest of this paper is organized as follows.

In section 2 it is defined, by three postulates that all tests should meet, what we mean when speaking of tests. In 2.1 some additional postulates are defined for tests in an AI environment. In section 3 we show how a general sort of test can be incorporated into the framework of [HLM93a]. In order to expose our ideas concerning tests in their purest and simplest form, it is initially assumed that the axiomatization for knowledge is given by the very weak system **KT**. In 3.2 it is shown how a different, more standard, axiomatization for knowledge, viz. **S5**, can be dealt with. In 3.3 it is shown that the general test action that we formalized in section 3 is indeed a test in the sense of section 2, and that the additional postulates for AI tests are also met. In section 4 we examine the relation between the properties that our test satisfy and the Gärdenfors postulates for belief expansion as defined in [Gär88]. It is shown that both our **KT** and **S5** implementations of tests meet the Gärdenfors postulates. In section 5 some related work concerning tests is discussed. Section 6 ends this paper with some conclusions and a guideline for further research.

2. The nature of tests

In this section we try and define the characteristic features of tests: what distinguishes actions that are tests from those that are not.

The most striking and probably most important feature of a test as we see it, is the fact that it is a means to acquire knowledge[2]. After an eye test is taken it is known whether the sight of the patient is satisfactory, after a driving test is taken it is known whether the candidate is actually fit to drive a car, etc. This acquisition of 'knowing whether' knowledge is of great importance in the theory of commonsense reasoning (cf. [Dav90], pp. 378–381, 386–388). The feature of tests that they provide knowledge to the person

---

[1]For an up-to-date report on the developments in this field see [HS93].

[2]The absence in dynamic logic of actions that aim solely at acquiring knowledge, and our conviction that the presence of such actions is highly desirable when modelling the behaviour of rational agents, was the main motivation underlying this research.

that performs the test concerning the truth or falsity of the proposition that is tested for, is referred to as *informativeness*, following [Moo84, MH].

Another desirable feature of tests as we see them, is the fact that the truth or falsity of the proposition that is tested for should not be affected by the actual act of testing, i.e., the outcome of the test should correspond to the situation in which the test is performed. In case of a driving test the actual ability of the candidate to drive a car should not be affected by the testing. Or consider the example of a dope test: if the fact that some athlete has taken drugs would be affected by the act of testing one could never be sure whether it is a case of doping. Note that in some real life situations it will not always be the case that the truth or falsity of the proposition that is tested, is invariant under the act of testing: it is for instance possible that due to the fact that the candidate is very nervous s/he will not pass her/his driving test, though s/he might be quite capable of driving safely. However, for our idealized tests, and in particular for the special AI tests that are described in 2.1, it seems to be an intuitively correct assumption to demand that the truth or falsity of the tested proposition is not affected by the actual act of testing. We refer to this second characteristic of tests as *truthfulness*.

The third characteristic feature of tests is closely connected with the second feature: not only should tests not affect the truth or falsity of the tested proposition, they should also not cause more changes to the world than absolutely necessary. In case of a driving test, the world is changed in the sense that the instructor has acquired knowledge concerning the candidate's driving abilities, the candidate has driven the car for half an hour, some time has elapsed, and possibly some other, related aspects, but it is for instance not the case that the world has become flat as a result of the driving test, or that Texas has become Mexican again. In other words, tests should behave as much as possible as *passive observations*. We refer to this third characteristic by stating that tests should meet the demand for *minimal change*.

2.1. Additional postulates for AI-testing

Besides the general demands that all tests should obey, there are some other postulates that are more or less specific for tests performed in an AI environment. In these environments usually only a small part of the world is formalized, and in this formalized part all laws of logic are strictly obeyed (the latter is not in general the case for real life situations). As such the demands for general tests as formulated above hold a fortiori when tests in an AI environment are considered.

An example of a typical AI situation is that of a rational, autonomous robot that is busy trying to accomplish some task, like for example building a tower out of blocks, or destroying ships (see example 3.9).

The following postulates provide additional demands that tests performed by rational agents in some AI environment should meet. The first two of these postulates are closely related to the basic demands given above.

2.1. DEFINITION. We state the following postulates concerning the behaviour of tests in AI environments:

(1) If the agent that performs the test has a priori knowledge on the truth or falsity of the proposition that is to be tested, then performing the test does not change anything. This postulate is related to informativeness and minimal change: if the information that is tested for is already available to the agent then the demand for minimal change ensures that nothing happens as a result of performing the test.

(2) Tests are idempotent, i.e., if two tests for the same proposition are performed sequentially, then the second test does not have any effect whatsoever. This postulate is related to informativeness and the first of the postulates given here: after the first test is performed, the agent has knowledge concerning truth or falsity of the proposition that was tested. This knowledge is a priori knowledge for the second test, which therefore does not change anything.

(3) The knowledge of the agent that performs the test concerning any aspects of the world should only be extended by testing. Should the agent for instance know that the block is green, then s/he still knows this after s/he has tested whether the block is lying on the table. This postulate provides a strengthening of minimal change: with regard to the a priori knowledge concerning aspects of the world, nothing may change as the result of performing a test.

(4) Testing for a 'positively phrased' proposition should be identical to a test for the 'negatively phrased' variant of the proposition, i.e, testing whether a block is lying on the table should be identical to testing whether the block is not lying on the table.

2.2. REMARK. All of the postulates given above, the general ones as well as the specific AI postulates, reflect the main emphasis of this paper: investigating properties of the results of testing. In particular do we not consider the actual procedures by means of which a test is performed.

As stated before, the postulates given in definition 2.1 provide natural demands when considering rational agents in an AI environment. Since the basic goal of our research is the formalization of the behaviour of this kind of rational agents (cf. [HLM93a, HLM93b]), we will ensure that the definitions for tests that we give in section 3 are such that not only the basic demands for tests (informativeness, truthfulness, minimal change), but also these additional postulates are met.

3. Formalizing tests in Kripke models

In this section we show how a general test action can be incorporated in a framework designed to formalize (part of) the behaviour of rational agents. This framework, defined in [HLM93a, HLM93b], considers the knowledge and abilities of agents, and the results of the actions that the agents may perform. To deal with the test action in a general, natural, and intuitively acceptable way, we propose to extend the framework of [HLM93a, HLM93b] by incorporating some aspects of *update semantics* (cf. [Vel91, EV93]). In particular, the parts from update semantics take care of a fairly natural and intuitively acceptable formalization of tests.

3.1. Epistemic updates in KT models

The models that we considered in [HLM93a, HLM93b] are tuples $< \mathcal{S}, \pi, \mathrm{R}, \mathbf{r}, \mathbf{c} >$, where $\mathcal{S}$ is a set of states, π is a valuation on propositional symbols, R formalizes the epistemic alternatives of agents, $\mathbf{r}$ deals with the state transitions caused by the performance of actions, and $\mathbf{c}$ deals with the agents' abilities. The actions that we considered in [HLM93a, HLM93b] were formalized as transitions between states in the models, this being completely in accordance with the usual approaches in dynamic logic (cf. [Har84, KT90, Gol92]). To deal with tests satisfactorily, we will no longer see actions as state transitions, but instead as transitions between pairs (model, state). For the class of actions that we originally considered, i.e, without the general test action, this new transition will come down to an ordinary state transition: the model is not modified, i.e, a transition from $(\mathcal{M}, s)$ to $(\mathcal{M}, s')$, where s' typically differs from s, results. However the test actions will in general affect the model under consideration whilst leaving the state as it is. This results in a transition from $(\mathcal{M}, s)$ to $(\mathcal{M}', s)$, where $\mathcal{M}'$ typically differs from $\mathcal{M}$.

As already explained in the previous section, tests are looked upon as means to acquire knowledge. Worded differently, tests are means to put an end to uncertainties of agents. For instance, should the agent be uncertain whether a block is lying on the table, after the appropriate test is performed the agent knows whether this is the case. As such this test reduces the number of epistemic alternatives that the agent has: all epistemic alternatives in which the truth value of the proposition 'the block is lying on the table' differs from the truth value in the current state, i.e., the state in which the test is performed, are no longer epistemically possible. The implementation of tests as we give it closely follows this intuitive idea: execution of a test action results in the removal of all epistemic alternatives that are no longer possible, i.e, a test performs an *epistemic update*.

3.1. EXAMPLE. Suppose that some agent currently resides in a world in which p holds and that the agent has two epistemic alternatives: one world in which p holds, one in which $\neg p$ holds. The result of the performance by the agent of the action **test** p is the removal of the epistemic alternative that differs from the current world in the truth value assigned to p. In a figure this epistemic update can be represented as follows (recall that the R-relation denotes epistemic alternatives) :
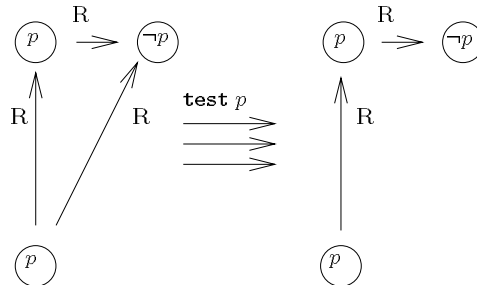

Figure 1: An epistemic update.

The test for p results in a new model in which the only epistemic alternative for the lowermost world is the one in which p holds.

In the formalization that follows some assumptions have been made. Firstly, we will allow only propositional formulae to be tested. The rationale underlying this assumption is the idea that only aspects of the real world can be observed by the agent: it is for instance not intuitive to allow agents to test for the knowledge of other agents since this would probably require communication rather than direct observations. Secondly, we will ensure that it is possible to express that some aspect of the real world cannot be observed by a certain agent. One can for instance think of a group of robots, each equipped with its own special sensor: one robot can test for radioactivity, the other for the intensity of sunlight, but it is not the case that one of them can test for both radioactivity and sunlight. These 'observation capacities' may be formalized through the agents' abilities.

3.1.1. The syntax

The language $\mathcal{L}$ used here is an extension of the language used in [HLM93a].

3.2. DEFINITION. The language $\mathcal{L}$ is based on a denumerable set Π of propositional symbols, a finite set At of atomic actions and a finite set $\mathcal{A} = \{1, \ldots, n\}$ of agents. The languages $\mathcal{L}$ and $\mathcal{L}_0$, where $\mathcal{L}_0$ is the purely propositional fragment of $\mathcal{L}$, are defined by the following BNF's:

$$\mathcal{L}_0: \ \psi ::= p \mid \neg\psi \mid \psi_1 \vee \psi_2$$

$$\mathcal{L}: \ \varphi ::= p \mid \neg\varphi \mid \varphi_1 \vee \varphi_2 \mid \mathbf{K}_i\varphi \mid {<}\mathrm{do}_i(\alpha){>}\varphi \mid \mathbf{A}_i\alpha$$

where α is from the class Ac given by:

$\alpha ::= a$		*atomic actions*
`confirm` φ		*confirmations*
$\alpha_1; \alpha_2$		*sequential composition*
`if` φ `then` α_1 `else` α_2 `fi`		*conditional composition*
`while` φ `do` α_1 `od`		*repetitive composition*
`test` ψ		<u>*test action*</u>

The language $\mathcal{L}_0$ is defined to be the set of all propositional formulae, i.e., $\mathcal{L}_0$ contains formulae ψ as given above.

The constructs $\wedge$, $\rightarrow$, and $\leftrightarrow$ are defined in the usual way. Other additional constructs are introduced by definitional abbreviation:

tt	is	$p \vee \neg p$
ff	is	$\neg\mathbf{tt}$
$[\mathrm{do}_i(\alpha)]\varphi$	is	$\neg {<}\mathrm{do}_i(\alpha){>}\neg\varphi$
`skip`	is	`confirm tt`
`fail`	is	`confirm ff`
α^0	is	`skip`
α^{n+1}	is	$\alpha; \alpha^n$

6

The action `confirm` φ behaves essentially like the test actions in dynamic logic (see [Har84, KT90, Gol92]). As such this action differs substantially from the genuine tests that we consider in this paper (see section 5).

The meaning of the other actions in Ac is respectively: the atomic action, sequential composition, conditional composition, and repetitive composition. Furthermore, `skip` denotes the empty action, and `fail` denotes the never succeeding action.

3.1.2. The semantics

As explained above, the Kripke models that we use are modified versions of the ones used in [HLM93a, HLM93b]. In order to make sure that the intuitive idea of 'not every agent is able to test everything' can be formalized in our framework, tests are treated as atomic actions with regard to abilities. In this way the 'observation capacities' of the agents can priori be defined.

3.3. DEFINITION. Let $\psi \in \mathcal{L}_0$. The equivalence class of ψ in propositional logic is defined by:

$$[\psi] = \{\psi' \in \mathcal{L}_0 \mid \models_{\mathrm{pl}} \psi \leftrightarrow \psi'\}$$

The set $[\psi, \neg \psi]$ is defined to be the union of $[\psi]$ and $[\neg \psi]$. The set $\mathrm{Eq}_{\mathcal{L}_0}$ is defined by

$$\mathrm{Eq}_{\mathcal{L}_0} = \{[\psi, \neg \psi] \mid \psi \in \mathcal{L}_0\}$$

3.4. DEFINITION. The class $\mathfrak{M}$ of Kripke models contains all tuples $\mathcal{M} = \langle \mathcal{S}, \pi, \mathrm{R}, \mathtt{r}, \mathtt{c} \rangle$ such that

(1) $\mathcal{S}$ is a set of possible worlds, or states.

(2) $\pi : \Pi \times \mathcal{S} \rightarrow \mathbf{bool}$ is a total function that assigns a truth value to propositional symbols in possible worlds.

(3) $\mathrm{R} : \mathcal{A} \rightarrow \wp(\mathcal{S} \times \mathcal{S})$ is a function that yields the epistemic accessibility relations for a given agent. For the sake of simplicity, we will initially consider knowledge to be axiomatized by $\mathbf{KT}$ (this according to the Chellas classification as given in [Che80]), i.e, the only demand on knowledge is that it should be veridical. This implies that $\mathrm{R}(i)$ is demanded to be a reflexive relation for all i. In section 3.2 we show how to deal with $\mathbf{S5}$ models, which are the models usually associated with knowledge, at least as far as the computer science literature is concerned. In these $\mathbf{S5}$ models, the $\mathrm{R}(i)$ relations are equivalence relations.

(4) $\mathtt{r} : \mathcal{A} \times At \rightarrow \mathcal{S} \rightarrow \wp(\mathcal{S})$ is such that $\mathtt{r}(i, a)(s)$ yields the (possibly empty) state transition in s caused by the event $\mathrm{do}_i(a)$. This function is such that for all atomic actions a it holds that $|\mathtt{r}(i,a)(s)| \leq 1$ for all i and s, where $|V|$ denotes the number of elements of the set V.

(5) $\mathtt{c} : \mathcal{A} \times At \cup \mathrm{Eq}_{\mathcal{L}_0} \rightarrow \mathcal{S} \rightarrow \mathbf{bool}$ is the capability function such that $\mathtt{c}(i,a)(s)$ indicates whether the agent i is capable of performing the action a in s, and $\mathtt{c}(i, [\psi, \neg\psi])(s)$ indicates whether the agent is able to test for ψ in the possible world s.

3.5. DEFINITION. Let $\mathcal{M} = < \mathcal{S}, \pi, \mathrm{R}, \mathbf{r}, \mathbf{c} >$ be some Kripke model. The set $\mathcal{M}_\subseteq$ of *epistemic sub-models* of $\mathcal{M}$ is defined by:

$$\mathcal{M}_\subseteq = \{\mathcal{M}' = < \mathcal{S}, \pi, \mathrm{R}', \mathbf{r}, \mathbf{c} > \in \mathfrak{M} \mid \forall i \in \mathcal{A}[\mathrm{R}'(i) \subseteq \mathrm{R}(i)]\}$$

3.6. DEFINITION. Let $\mathcal{M} = < \mathcal{S}, \pi, \mathrm{R}, \mathbf{r}, \mathbf{c} >$ be some Kripke model from $\mathfrak{M}$. For propositional symbols, negated formulae, conjunctions, and epistemic formulae, $\mathcal{M}, s \models \varphi$ is inductively defined as usual. For the other clauses $\mathcal{M}, s \models \varphi$ is defined as follows:

$$\mathcal{M}, s \models <\mathrm{do}_i(\alpha)>\varphi \quad\Leftrightarrow \exists \mathcal{M}', s' \in \mathcal{M}_\subseteq \times \mathcal{S}[\mathcal{M}', s' \in \mathbf{r}(i, \alpha)(\mathcal{M}, s)$$
$$\& \, \mathcal{M}', s' \models \varphi]$$
$$\mathcal{M}, s \models \mathbf{A}_i \alpha \quad\Leftrightarrow \mathbf{c}(i, \alpha)(\mathcal{M}, s) = \mathbf{1}$$

where $\mathbf{r}$ and $\mathbf{c}$ are defined by:

$\mathbf{r}$	$: (\mathcal{A} \times Ac) \to ((\mathcal{M}_\subseteq \times \mathcal{S}) \cup \mathcal{S}) \to \wp(\mathcal{M}_\subseteq \times \mathcal{S})$
$\mathbf{r}(i, a)(\mathcal{M}, s)$	$= \mathcal{M}, \mathbf{r}(i, a)(s)$
$\mathbf{r}(i, \mathtt{confirm}\ \varphi)(\mathcal{M}, s)$	$= \{(\mathcal{M}, s)\}$ if $\mathcal{M}, s \models \varphi$
	$= \emptyset$ otherwise
$\mathbf{r}(i, \alpha_1; \alpha_2)(\mathcal{M}, s)$	$= \mathbf{r}(i, \alpha_2)(\mathbf{r}(i, \alpha_1)(\mathcal{M}, s))$
$\mathbf{r}(i, \mathtt{if}\ \varphi\ \mathtt{then}\ \alpha_1\ \mathtt{else}\ \alpha_2\ \mathtt{fi})(\mathcal{M}, s)$	$= \mathbf{r}(i, \alpha_1)(\mathcal{M}, s)$ if $\mathcal{M}, s \models \varphi$
	$= \mathbf{r}(i, \alpha_2)(\mathcal{M}, s)$ otherwise
$\mathbf{r}(i, \mathtt{while}\ \varphi\ \mathtt{do}\ \alpha_1\ \mathtt{od})(\mathcal{M}, s)$	$= \{(\mathcal{M}', s') \in \mathcal{M}_\subseteq \times \mathcal{S} \mid \exists k \in \mathbb{N}$

$$\exists \mathcal{M}_0, s_0 \ldots \exists \mathcal{M}_k, s_k \in \mathcal{M}_\subseteq \times \mathcal{S}$$
$$[\mathcal{M}_0, s_0 = \mathcal{M}, s \,\&\, \mathcal{M}_k, s_k = \mathcal{M}', s' \,\&\, \forall j < k$$
$$[\mathcal{M}_{j+1}, s_{j+1} \in \mathbf{r}(i, \mathtt{confirm}\ \varphi; \alpha_1)(\mathcal{M}_j, s_j)] \,\&$$
$$\mathcal{M}', s' \models \neg\varphi]\}$$

and

$\mathbf{r}(i, \mathtt{test}\ \varphi)(\mathcal{M}, s)$	$= (\mathrm{Eu}_\varphi(\mathcal{M}, s), s)$, where $\mathcal{M} = < \mathcal{S}, \pi, \mathrm{R}, \mathbf{r}, \mathbf{c} >$,

and $\mathrm{Eu}_\varphi(\mathcal{M}, s) = < \mathcal{S}, \pi, \mathrm{R}', \mathbf{r}, \mathbf{c}' > \in \mathcal{M}_\subseteq$, with
$\mathrm{R}'(j) = \mathrm{R}(j)$ for $j \neq i$, and
$\mathrm{R}'(i) = \mathrm{R}(i) \setminus \{(s, s') \mid \mathrm{not}(\mathcal{M}, s \models \varphi \Leftrightarrow \mathcal{M}, s' \models \varphi)\}$
where $\mathbf{r}(i, \alpha)(\emptyset) \qquad = \emptyset$
and

$\mathbf{c}$	$: \mathcal{A} \times (Ac \cup \mathrm{Eq}_{\mathcal{L}_0}) \to ((\mathcal{M}_\subseteq \times \mathcal{S}) \cup \mathcal{S}) \to \mathbf{bool}$
$\mathbf{c}(i, a)(\mathcal{M}, s)$	$= \mathbf{c}(i, a)(s)$
$\mathbf{c}(i, \mathtt{confirm}\ \varphi)(\mathcal{M}, s)$	$= \mathbf{1}$ if $\mathcal{M}, s \models \varphi$
	$= \mathbf{0}$ otherwise
$\mathbf{c}(i, \alpha_1; \alpha_2)(\mathcal{M}, s)$	$= \mathbf{c}(i, \alpha_1)(\mathcal{M}, s) \,\&\, \mathbf{c}(i, \alpha_2)(\mathbf{r}(i, \alpha_1)(\mathcal{M}, s))$
$\mathbf{c}(i, \mathtt{if}\ \varphi\ \mathtt{then}\ \alpha_1\ \mathtt{else}\ \alpha_2\ \mathtt{fi})(\mathcal{M}, s)$	$= \mathbf{c}(i, \mathtt{confirm}\ \varphi; \alpha_1)(\mathcal{M}, s)$ or
	$\mathbf{c}(i, \mathtt{confirm}\ \neg\varphi; \alpha_2)(\mathcal{M}, s)$
$\mathbf{c}(i, \mathtt{while}\ \varphi\ \mathtt{do}\ \alpha_1\ \mathtt{od})(\mathcal{M}, s)$	$= \mathbf{1}$ if $\exists k \in \mathbb{N}[\mathbf{c}(i, (\mathtt{confirm}\ \varphi; \alpha_1)^k;$

$$\mathtt{confirm}\ \neg\varphi)(\mathcal{M}, s) = 1]$$

	$= \mathbf{0}$ otherwise
$\mathbf{c}(i, \mathtt{test}\ \psi)(\mathcal{M}, s)$	$= \mathbf{c}(i, [\psi, \neg\psi])(s)$
where $\mathbf{c}(i, \alpha)(\emptyset)$	$= \mathbf{1}$.

The notions of satisfiability and validity are defined as usual.

When agent i performs a test for φ in state s, Eu_φ removes from the set $\mathrm{R}(i)$ that denotes the epistemic accessibility relation for agent i all those pairs (s, s') in which the truth value assigned to φ in s' differs from that assigned in s. In this way the *epistemic update with φ*, as it results after a test for φ is performed, is nicely formalized by Eu_φ. The definition of the function r for the other actions is the standard one for the class of deterministic while programs (see for instance [HR83]).

As explained at the beginning of 3.1.2, tests are treated as atomic actions with regard to abilities, this in order to allow for optimal flexibility concerning the observations that agents are able to make. With regard to the abilities for the other actions, the motivation for the choices made in definition 3.6 is the following. The definition of $\mathrm{c}(i, \mathtt{confirm}\, \varphi)(s)$ is based on the idea that this action is one that searches for confirmation of φ; if φ does not hold, it is not possible to get confirmation for φ, and hence the agent is not capable of confirming φ. As such the behaviour of the action $\mathtt{confirm}\, \varphi$ is essentially identical to that of the dynamic logic 'test' as it is usually defined (see also section 5). An agent is capable of performing a sequential composition $\alpha_1; \alpha_2$ iff s/he is capable of α_1, and s/he is capable of performing α_2 after s/he has performed α_1. An agent is capable of performing a conditional composition, if s/he is able to confirm either the condition or the negation of the condition, and thereafter s/he is capable of performing the then-part or the else-part, dependent on whether the condition or the negation of the condition is confirmed. Lastly, an agent is capable of performing a repetitive composition $\mathtt{while}\, \varphi\, \mathtt{do}\, \alpha_1\, \mathtt{od}$ iff s/he is able to perform the action $(\mathtt{confirm}\, \varphi; \alpha_1)^k; \mathtt{confirm}\, \neg\varphi$ for some $k \in \mathbb{N}$. Note that the definitions for both the conditional and the repetitive composition are based on the defining equalities that hold in dynamic logic (see [Har84, KT90, Gol92]).

The implementation of tests by means of epistemic updates is sound to the extent that **KT** models are transformed into **KT** models.

3.7. LEMMA. *Let $\mathcal{M} = <\mathcal{S}, \pi, \mathrm{R}, \mathrm{r}, \mathrm{c}>$ be some* **KT** *model, let $s \in \mathcal{S}$, and let $i \in \mathcal{A}$ and $\varphi \in \mathcal{L}_0$. Now $\mathcal{M}'$ such that $\mathrm{r}(i, \mathtt{confirm}\, \varphi)(\mathcal{M}, s) = \mathcal{M}', s$ is a* **KT** *model.*

3.8. REMARK. Besides treating tests as atomic actions with regard to abilities, some alternative approaches can be thought of. It is for instance possible to interpret abilities for the $\mathtt{test}$ action as *inclinations*, i.e, instead of stating that an agent is able to test for ψ it is stated that s/he is inclined to do so. A natural way to model these inclinations is given by assuming that agents are inclined to test for some subject if they do not have any information concerning this subject on beforehand. This assumption could lead to the following definition of $\mathrm{c}(i, \mathtt{test}\, \varphi)(\mathcal{M}, s)$:

$$\mathrm{c}(i, \mathtt{test}\, \varphi)(\mathcal{M}, s) = 1 \Leftrightarrow \mathcal{M}, s \models \neg \mathrm{K}_i \varphi \wedge \neg \mathrm{K}_i \neg \varphi$$

When interpreting abilities for tests as inclinations it is possible to model stronger degrees of informativeness, such as those given in definition 3.19.

The following example shows how tests as we defined them can be used in a typical AI application: the formalization of the behaviour of some rational, autonomous robot, in this case an automated submarine.

9

3.9. Example (The little submarine). Some little automated submarine is making a trip under the surface of the sea. It is programmed to fire a torpedo as soon as it is sure that a ship is around. Every time it has covered another sea mile, the submarine tests whether there is a ship around by putting up its periscope. Just before the periscope is put up, the submarine is not sure whether there actually is a ship around: it considers both a world in which there is one around, and a world in which there is not, as epistemically possible. After the periscope has been put up the submarine knows which of its epistemic alternatives needs to be removed. Figure 2 shows schematically how a trip of the submarine could look like.
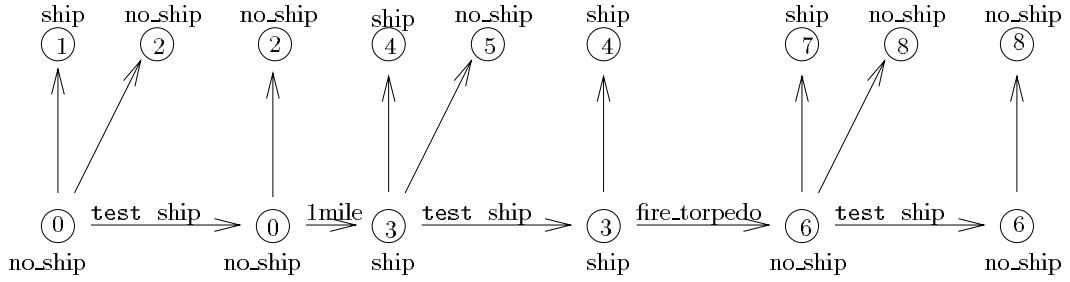

Figure 2: A trip of the little submarine

One should pay attention to the fact that a test only modifies the model under consideration, and does not cause a state transition: only the epistemic state is changed, not the state of the world. For that reason the model that corresponds to the initial situation sketched above contains only nine different states, corresponding to the numbers 0 to 8 in figure 2. This initial model $\mathcal{M} = < \mathcal{S}, \pi, \mathrm{R}, \mathrm{r}, \mathrm{c} >$ is defined as given below. In this definition sm is used to abbreviate agent 'submarine', b stands for the proposition 'ship (boat) around', γ is the action of going ahead for 1 mile, and τ is the action of firing a torpedo.

(1) $\mathcal{S} = \{s_0, \ldots, s_8\}$,

(2) $\pi(b, s_j) = \mathbf{1}$ iff $j \in \{1, 3, 4, 7\}$,
$\pi(b, s_j) = \mathbf{0}$ iff $j \in \{0, 2, 5, 6, 8\}$.

(3) $\mathrm{R}(sm) = \{(s_0, s_1), (s_0, s_2), (s_3, s_4), (s_3, s_5), (s_6, s_7), (s_6, s_8)\} \cup \{(s_j, s_j) \mid j \in \{0, \ldots, 8\}\}$.

(4) $\mathrm{r}(sm, \gamma)(s_0) = s_3$, $\mathrm{r}(sm, \tau)(s_3) = s_6$.

(5) c is arbitrary.

This model indeed formalizes the behaviour of the little submarine in an intuitively acceptable way:

3.10. Proposition. *The following are true in the model $\mathcal{M}$:*

- $\mathcal{M}, s_0 \models < \mathrm{do}_{sm}(\texttt{test } b) > \mathbf{K}_{sm} \neg b$, *i.e., after putting up the periscope, the submarine knows that there is no ship around.*

- $\mathcal{M}, s_0 \models < \mathrm{do}_{sm}(\gamma) > \neg \mathbf{K}_{sm} b \wedge \neg \mathbf{K}_{sm} \neg b$, *i.e., after having been under the surface for one mile, the submarine is not sure whether there is some ship around.*

- $\mathcal{M}, s_0 \models < \mathrm{do}_{sm}(\gamma; \texttt{test } b) > \mathbf{K}_{sm} b$, *i.e., by putting up its periscope after the one mile under sea, the submarine knows that there is some ship around.*

10

- $\mathcal{M}, s_3 \models <\mathrm{do}_{sm}(\tau; \texttt{test}\ b) > \mathbf{K}_{sm}\neg b$, *i.e, by putting up its periscope after the torpedo is fired, the submarine knows that the ship is destroyed.*

Note in particular that the putting up of the periscope and the knowledge acquired by doing this, is modelled in a very nice way using our `test` actions. Note furthermore that the action γ is in some sense *anti-informative*: by performing this action the (objective) knowledge of the submarine decreases.

Figure 3 depicts schematically how the knowledge of the submarine changes as a result of its testing.

(Model, state) transition	*Knowledge before testing:*	*Knowledge after testing:*
$\mathcal{M}, s_0 \mapsto \mathcal{M}', s_0$	$\neg\mathbf{K}_{sm}b \wedge \neg\mathbf{K}_{sm}\neg b$	$\mathbf{K}_{sm}\neg b$
$\mathcal{M}', s_3 \mapsto \mathcal{M}'', s_3$	$\neg\mathbf{K}_{sm}b \wedge \neg\mathbf{K}_{sm}\neg b$	$\mathbf{K}_{sm}b$
$\mathcal{M}'', s_6 \mapsto \mathcal{M}''', s_6$	$\neg\mathbf{K}_{sm}b \wedge \neg\mathbf{K}_{sm}\neg b$	$\mathbf{K}_{sm}\neg b$

Figure 3: The knowledge of the submarine.

In figure 3, $\mathcal{M}'$ is the model that is the result of performing the test for the proposition 'ship around' in the state s_0, $\mathcal{M}''$ and $\mathcal{M}'''$ are the results of performing the same test in the states s_3 and s_6 respectively.

The following example highlights some characteristic features of definition 3.6.

3.11. EXAMPLE. Consider a Kripke model that contains 5 possible worlds such that the function π is as given in figure 4, and the relation R is for agent i the reflexive closure of the arcs given in figure 4.
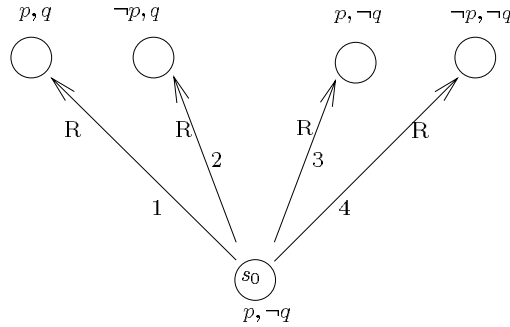

Figure 4: The valuation and epistemic accessibility relation.

The result of agent i in s_0 performing a test for 10 different formulae containing p and q is as given in figure 5. In this table 'remove j' means that the pair corresponding to the arc labeled with j is removed from R(i).

Note that in this model it is indeed the case that testing for a formula and testing for the negation of the formula have the same effect (cf. the fourth of the postulates of section 2.1). Note furthermore that if two formulae φ and ψ are equivalent in propositional logic, a test for φ causes the same epistemic update as a test for ψ. For instance `test` $(p \rightarrow q)$ leads to the removal of arcs 1, 2, and 4 (note that $(p \rightarrow q)$ is equivalent to $(\neg p \vee q)$), and

Test for:	Result	Test for:	Result
p	remove $2, 4$	q	remove 1,2
$p \vee q$	remove 4	$\neg p \wedge \neg q$	remove 4
$p \vee \neg q$	remove 2	$\neg p \wedge q$	remove 2
$\neg p \vee q$	remove 1,2,4	$p \wedge \neg q$	remove 1,2,4
$\neg p \vee \neg q$	remove 1	$p \wedge q$	remove 4

Figure 5: The results of testing.

test $(q \rightarrow p)$ results in arc 2 being removed, the same arc that is removed when a test for the (equivalent) formula $(\neg q \vee p)$ is performed.

Another interesting point that this example shows is that successively testing for two formulae is not equivalent to testing for the conjunction of the formulae. The result of performing test p; test q is that the arcs 1,2, and 4 are removed, and as a consequence the agent knows $p \wedge \neg q$. Performing test $(p \wedge q)$ results in a removal of the arc 1 and yields only $(\neg p \vee \neg q)$ as knowledge for the agent. However, note that test p; test $\neg q$ and test $(p \wedge \neg q)$ do have identical results (this latter property is formalized in lemma 3.36).

The observations made concerning the difference between successively testing for two formulae and testing for the conjunction of those formulae are intuitively well founded. Consider for instance the situation of an agent that is testing whether s/he is smart and pretty by asking the magical mirror whether this is the case. If the mirror says 'no', the agent only knows that s/he is not smart and pretty. Should s/he ask first whether s/he is smart ('no'), and thereafter whether s/he is pretty ('yes'), the agent does both know that s/he is not smart, and that s/he is pretty, which comprises more information than just knowing that s/he is not smart and pretty.

An important remark that needs to be made is that when allowing the possibility to test for non-propositional formulae, it is not all too clear that the demand for truthfulness is a meaningful one. Since tests are means to acquire knowledge, and epistemic formulae ψ represent the knowledge of agents, it is intuitively clear that the truth of ψ can be affected as a result of testing; this is exactly what happens in example 3.12.

3.12. EXAMPLE.[3] Consider a Kripke model $\mathcal{M} = < \mathcal{S}, \pi, \mathrm{R}, \mathtt{r}, \mathtt{c} >$ where

(1) $\mathcal{S} = \{s_0, s_1, s_2\}$,

(2) $\pi(p, s_j) = \mathbf{0}$ iff $j = 2$,

(3) $\mathrm{R}(i) = \{(s_0, s_1), (s_0, s_2), (s_1, s_2)\} \cup \{(s_j, s_j) \mid j = 0, 1, 2\}$,

(4) $\mathtt{r}$ and $\mathtt{c}$ are arbitrary.

Suppose that the agent i performs a test for the formula $\varphi \stackrel{\text{def}}{=} p \rightarrow \mathbf{K}_i p$ in state s_0. The formula φ holds in s_2 and does not hold in s_0 and s_1. Therefore after the test is performed, s_2 no longer is an epistemic alternative for the agent in s_0, i.e., a model $\mathcal{M}'$ results such that $\mathrm{R}'(i) = \mathrm{R}(i) \setminus \{(s_0, s_2)\}$. For $\mathcal{M}'$ it holds that

$$\mathcal{M}', s_0 \models \varphi$$

---

[3]This example is due to Frans Voorbraak.

and since $\mathcal{M}, s_0 \not\models \varphi$ this test does not satisfy the demand for truthfulness, i.e, the truth value assigned to φ in s is changed as a result of testing for φ.

As already stated at the beginning of this section, it seems intuitively correct to allow testing for propositional formulae only. For this reason the problems that occur in example 3.12 with tests for epistemic formulae are of no real concern to us, and are certainly no reason to discard the formalization presented here.

3.2. Epistemic updates in S5 models

When formalizing tests as we did in definition 3.6, it was assumed that the knowledge of agents is modelled by a **KT** axiomatization. However, in [HLM93a, HLM93b] we formalized knowledge using an **S5** axiomatization, this in accordance with the common practice in computer science (cf. [HM85, KL88, MH]).

In this section we show how the definition of the epistemic update function Eu_φ as given in 3.6 needs to be modified in order to adequately deal with **S5** models. The problem with epistemic updates in these models is that it is not possible to straightforwardly remove arcs that no longer fit in: the resulting model should be an **S5** one, which means that the $R(i)$ relations need to be equivalence relations.

Figure 6 shows how the straightforward implementation of epistemic updates as given in definition 3.6 affects an **S5** model. When the agent i performs a `test` p action in the state s_0, the resulting model no longer satisfies the demands for **S5**. This can for instance be seen since it holds that $\mathcal{M}, s_0 \models < \text{do}_i(\texttt{test } p) > (\mathbf{K}_i p \wedge \neg \mathbf{K}_i \mathbf{K}_i p)$, i.e, in the model that results after the test is performed, the $R(i)$ relation no longer is transitive, and therefore this model certainly is not an **S5** model.
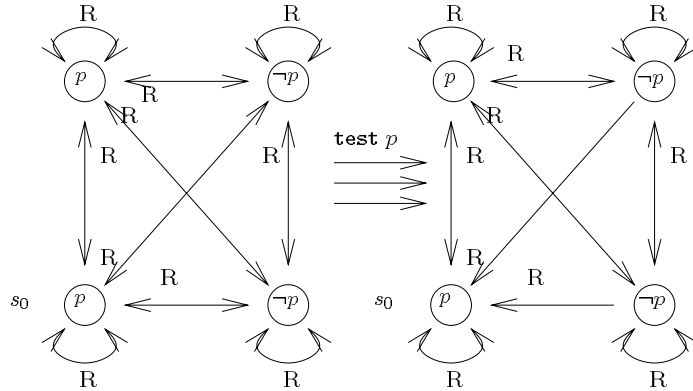

Figure 6: Updating an **S5** model in a **KT** style.

Our approach towards epistemic updates in **S5** models is based on the idea that the possible worlds that contradict the newly acquired knowledge are expelled from the epistemic equivalence class they belonged to before the test was performed. As a result of performing a test for φ the original epistemic equivalence class of a given state s is divided into two new equivalence classes: one that contains all states that agree with s on the truth value
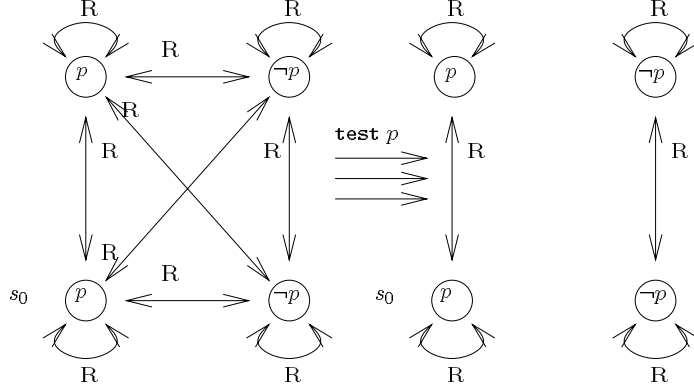
13

Figure 7: An epistemic update in an **S5** model.

of φ, and one that contains all states that do not agree. In figure 7 it is shown what this means for the implementation of the epistemic update.

An additional definition is necessary before the modified epistemic update function Eu_φ can be given.

3.13. DEFINITION. Let $\mathcal{M} =< \mathcal{S}, \pi, \mathrm{R}, \mathbf{r}, \mathbf{c} >$ be some **S5** model, let $s \in \mathcal{S}$, let $\mathcal{S}' \subseteq \mathcal{S}$, and let $\varphi \in \mathcal{L}_0$.

- $[s]_{\mathrm{R}(i)} \overset{\mathrm{def}}{=} \{s' \in \mathcal{S} \mid (s, s') \in \mathrm{R}(i)\}$.
- $[s]_{\mathrm{R}(i)}^{\varphi+} \overset{\mathrm{def}}{=} [s]_{\mathrm{R}(i)} \cap \{s' \in \mathcal{S} \mid \mathcal{M}, s \models \varphi \Leftrightarrow \mathcal{M}, s' \models \varphi\}$.
- $[s]_{\mathrm{R}(i)}^{\varphi-} \overset{\mathrm{def}}{=} [s]_{\mathrm{R}(i)} \cap \{s' \in \mathcal{S} \mid \mathrm{not}(\mathcal{M}, s \models \varphi \Leftrightarrow \mathcal{M}, s' \models \varphi)\} = [s]_{\mathrm{R}(i)} \setminus [s]_{\mathrm{R}(i)}^{\varphi+}$.
- $\mathrm{Cl}_{\mathrm{eq}}(\mathcal{S}') \overset{\mathrm{def}}{=} \mathcal{S}' \times \mathcal{S}'$.

3.14. REMARK. The set $[s]_{\mathrm{R}(i)}$ is the $\mathrm{R}(i)$ equivalence class of s. It contains all states that are in the relation $\mathrm{R}(i)$ with s. The set $[s]_{\mathrm{R}(i)}^{\varphi+}$ contains all elements from $[s]_{\mathrm{R}(i)}$ that agree with s on the truth value assigned to φ. Analogously, $[s]_{\mathrm{R}(i)}^{\varphi-}$ is the set that contains all elements from $[s]_{\mathrm{R}(i)}$ that do not agree with s on the truth value assigned to φ. The function $\mathrm{Cl}_{\mathrm{eq}}$ denotes the closure under equivalence of a certain set of possible worlds: it simply states that all possible worlds in the set are pairwise related.

Our new function Eu_φ formalizes the idea that as a result of performing a test for φ the original epistemic equivalence class of s is divided into two new equivalence classes: one that contains all states that agree with s on the truth value of φ, i.e, the states from $[s]_{\mathrm{R}(i)}^{\varphi+}$, and one that contains all states that do not agree, these are the states from $[s]_{\mathrm{R}(i)}^{\varphi-}$.

3.15. DEFINITION (Epistemic updates, **S5** style). Let $\mathcal{M} =< \mathcal{S}, \pi, \mathrm{R}, \mathbf{r}, \mathbf{c} >$ be some **S5** model. The function $\mathrm{Eu}_\varphi(\mathcal{M}, s)$ is defined by:

$\mathrm{Eu}_\varphi(\mathcal{M}, s) =< \mathcal{S}', \pi', \mathrm{R}', \mathbf{r}', \mathbf{c}' >$ where
$\quad \mathcal{S}' = \mathcal{S}, \pi' = \pi, \mathbf{r}' = \mathbf{r}, \mathbf{c}' = \mathbf{c}$ and
$\quad \mathrm{R}'(j) = \mathrm{R}(j)$ for $j \neq i$,
$\quad \mathrm{R}'(i) = (\mathrm{R}(i) \setminus \mathrm{Cl}_{\mathrm{eq}}([s]_{\mathrm{R}(i)})) \cup \mathrm{Cl}_{\mathrm{eq}}([s]_{\mathrm{R}(i)}^{\varphi+}) \cup \mathrm{Cl}_{\mathrm{eq}}([s]_{\mathrm{R}(i)}^{\varphi-})$

14

3.16. REMARK. The epistemic updates **S5** style could be seen as consisting of a sequence of two actions: first the arrows are removed in an identical way as with a **KT** update, and thereafter a reshaping takes place such that an **S5** model results. A particular consequence of this implementation is that after the test, the number of epistemic alternatives of the state in which the test is performed is as large as possible given the demand of informativeness.

Using this modified definition of the Eu_φ function it is indeed the case that **S5** models are transformed into **S5** models:

3.17. LEMMA. *Let $\mathcal{M} = < \mathcal{S}, \pi, \text{R}, \text{r}, \text{c} >$ be some **S5** model and let $s \in \mathcal{S}$, $i \in \mathcal{A}$ and $\varphi \in \mathcal{L}_0$. Now $\mathcal{M}'$ such that $\text{r}(i, \text{confirm } \varphi)(\mathcal{M}, s) = \mathcal{M}', s$ is an **S5** model.*

The differences between tests in an **S5** setting and tests in a **KT** setting are elaborated on in section 3.3.

3.3. Properties of the `test` action

In this section we consider some properties of the tests formalized in sections 3.1 and 3.2. Furthermore some results are given concerning the differences between epistemic updates in a **KT** setting and updates **S5** style.

We start by defining the formal counterparts of the notions of informativeness and truthfulness. These definitions are given for general actions, and not specifically for the `test` action of section 3.

As already indicated in section 2, an action is informative with regard to a certain proposition if the actor that executes the has afterwards test either knowledge that the proposition indeed holds, or that the proposition does not hold. Following [MH], we define informativeness as follows.

3.18. DEFINITION. Let $\mathcal{M} = < \mathcal{S}, \pi, \text{R}, \text{r}, \text{c} >$ be some Kripke model, and let $s \in \mathcal{S}$. The event $\text{do}_i(\alpha)$ is informative in s with regard to formula $\varphi \in \mathcal{L}$ if and only if

$$\mathcal{M}, s \models [\text{do}_i(\alpha)](\mathbf{K}_i\varphi \vee \mathbf{K}_i\neg\varphi)$$

holds. The event $\text{do}_i(\alpha)$ is informative in $\mathcal{M}$ with regard to φ iff it is informative in all $s \in \mathcal{S}$.

The following definition introduces some stronger notions of informativeness. The first of these notions (strict informativeness) is suggested in [MH].

3.19. DEFINITION. Let $\mathcal{M} = < \mathcal{S}, \pi, \text{R}, \text{r}, \text{c} >$ be some Kripke model, and let $s \in \mathcal{S}$.

(1) The event $\text{do}_i(\alpha)$ is *strictly informative* in s with regard to formula $\varphi \in \mathcal{L}$ if and only if $\text{do}_i(\alpha)$ is informative with regard to φ in s and furthermore

$$\mathcal{M}, s \models (\neg\mathbf{K}_i\varphi \wedge \neg\mathbf{K}_i\neg\varphi)$$

holds.

(2) The event $\mathrm{do}_i(\alpha)$ is *highly informative* in s with regard to formula $\varphi \in \mathcal{L}$ if and only if $\mathrm{do}_i(\alpha)$ is strictly informative with regard to φ in s and furthermore

$$\mathcal{M}, s \models \neg \mathbf{K}_i[\mathrm{do}_i(\alpha)]\varphi \wedge \neg \mathbf{K}_i[\mathrm{do}_i(\alpha)]\neg\varphi$$

holds.

(3) The event $\mathrm{do}_i(\alpha)$ is *strongly informative* in s with regard to formula $\varphi \in \mathcal{L}$ if and only if $\mathrm{do}_i(\alpha)$ is informative with regard to φ in s and furthermore

$$\mathcal{M}, s \models <\mathrm{do}_i(\alpha)>(\mathbf{K}_i\varphi \vee \mathbf{K}_i\neg\varphi)$$

holds.

3.20. REMARK. Note that due to the determinacy of our framework the definition of strong informativeness could be simplified to the equivalent formula $<\mathrm{do}_i(\alpha)>(\mathbf{K}_i\varphi \vee \mathbf{K}_i\neg\varphi)$. Note furthermore that the following implications hold by definition:

- highly informative $\Rightarrow$ strictly informative $\Rightarrow$ informative.
- strongly informative $\Rightarrow$ informative.

The demand for truthfulness is even easier to formalize. It simply states that the truth value of the proposition that is to be tested, is the same before and after the test.

3.21. DEFINITION. Let $\mathcal{M} =< \mathcal{S}, \pi, \mathrm{R}, \mathbf{r}, \mathbf{c} >$ be some Kripke model, and let $s \in \mathcal{S}$. The event $\mathrm{do}_i(\alpha)$ is truthful in s with regard to formula $\varphi \in \mathcal{L}$ if and only if

$$\mathcal{M}, s \models (\varphi \rightarrow [\mathrm{do}_i(\alpha)]\varphi) \wedge (\neg\varphi \rightarrow [\mathrm{do}_i(\alpha)]\neg\varphi)$$

holds. The event $\mathrm{do}_i(\alpha)$ is truthful in $\mathcal{M}$ with regard to φ iff it is truthful in all $s \in \mathcal{S}$.

Note that definition 3.21 implies that events $\mathrm{do}_i(\alpha)$ that are impossible in a given world $\mathcal{M}, s$, i.e, $\mathcal{M}, s \not\models <\mathrm{do}_i(\alpha)> \mathbf{tt}$, are trivially truthful in that world. This is exactly the reason why the actions that are considered to be tests in standard dynamic logic (see [Har84, KT90, Gol92]) are truthful (see section 5). A natural strengthening of definition 3.21 is the following.

3.22. DEFINITION. Let $\mathcal{M} =< \mathcal{S}, \pi, \mathrm{R}, \mathbf{r}, \mathbf{c} >$ be some Kripke model, and let $s \in \mathcal{S}$. The event $\mathrm{do}_i(\alpha)$ is *strongly truthful* in s with regard to formula $\varphi \in \mathcal{L}$ if and only if $\mathrm{do}_i(\alpha)$ is truthful with regard to φ and furthermore

$$\mathcal{M}, s \models (\varphi \rightarrow <\mathrm{do}_i(\alpha)> \varphi) \wedge (\neg\varphi \rightarrow <\mathrm{do}_i(\alpha)> \neg\varphi)$$

holds. The event $\mathrm{do}_i(\alpha)$ is strongly truthful in $\mathcal{M}$ with regard to φ iff it is strongly truthful in all $s \in \mathcal{S}$.

3.23. REMARK. Note that, again due to the determinacy of our framework, the definition of strong truthfulness could be simplified to the equivalent formula $(\varphi \rightarrow <\mathrm{do}_i(\alpha)> \varphi) \wedge (\neg\varphi \rightarrow <\mathrm{do}_i(\alpha)> \neg\varphi)$, i.e, truth of the latter formula implies truthfulness of $\mathrm{do}_i(\alpha)$ with regard to φ.

As formalized in lemma 3.25, our test actions are strongly truthful, this in contrast with the 'tests' of dynamic logic (see section 5).

The demand for minimal change is not as straightforwardly formalized as informativeness and truthfulness. Instead of a formal definition we will use the following intuitive one.

3.24. DEFINITION. An event satisfies the demand for minimal change if the end pair $(\mathcal{M}', s')$ of the event resembles the begin pair $(\mathcal{M}, s)$ *as closely as possible.* This resemblance is to be as complete as possible, i.e, not only should the worlds s' and s be similar (preferably identical), but also the R, π, r and c functions should be taken into account for all agents.

Using the formalizations of informativeness and truthfulness as given above, it turns out that both in a **KT** and in an **S5** setting the test action as we defined it indeed satisfies these first two demands. In lemma 3.25, $\models$ denotes both validity in the class of **KT** models and in the class of **S5** models.

3.25. LEMMA. *For all agents i, and for all formulae $\varphi, \psi \in \mathcal{L}_0$ we have:*

- $\models [\mathrm{do}_i(\texttt{test } \varphi)]\mathbf{K}_i\varphi \vee \mathbf{K}_i\neg\varphi$, *i.e., tests are indeed informative.*
- $\models <\mathrm{do}_i(\texttt{test } \varphi)>\mathbf{K}_i\varphi \vee \mathbf{K}_i\neg\varphi$, *i.e, tests are even strongly informative.*
- $\models (\varphi \rightarrow [\mathrm{do}_i(\texttt{test } \varphi)]\varphi) \wedge (\neg\varphi \rightarrow [\mathrm{do}_i(\texttt{test } \varphi)]\neg\varphi)$, *i.e., tests meet the demand for truthfulness.*
- $\models (\varphi \rightarrow <\mathrm{do}_i(\texttt{test } \varphi)>\varphi) \wedge (\neg\varphi \rightarrow <\mathrm{do}_i(\texttt{test } \varphi)>\neg\varphi)$, *i.e., tests are strongly truthful.*
- $\models \mathbf{K}_i\psi \rightarrow [\mathrm{do}_i(\texttt{test } \varphi)]\mathbf{K}_i\psi$, *i.e, the knowledge of the agent that performs the test concerning any aspects of the world is only extended by testing.*

3.26. LEMMA. *For all agents i, and for all formulae $\varphi \in \mathcal{L}_0$ we have:*

- $\models (\varphi \leftrightarrow [\mathrm{do}_i(\texttt{test } \varphi)]\varphi) \wedge ([\mathrm{do}_i(\texttt{test } \varphi)]\varphi \leftrightarrow <\mathrm{do}_i(\texttt{test } \varphi)>\varphi)$.
- $\models \neg\mathbf{K}_i\varphi \leftrightarrow \neg\mathbf{K}_i([\mathrm{do}_i(\texttt{test } \varphi)]\varphi)$, *which implies that for our* test *actions, strict and high informativeness are equivalent notions.*
- $\models (\varphi \rightarrow (<\mathrm{do}_i(\texttt{test } \varphi)>\mathbf{K}_i\varphi \wedge [\mathrm{do}_i(\texttt{test } \varphi)]\mathbf{K}_i\varphi)) \wedge$
  $(\neg\varphi \rightarrow (<\mathrm{do}_i(\texttt{test } \varphi)>\mathbf{K}_i\neg\varphi \wedge [\mathrm{do}_i(\texttt{test } \varphi)]\mathbf{K}_i\neg\varphi))$.

3.27. REMARK. Two important remarks need to be made concerning the last item of lemma 3.25. Firstly, a test for a formula φ may result in the acquisition of knowledge concerning some totally unrelated formula ψ. Secondly, with regard to epistemic formula it is possible that tests result in a genuine knowledge revision (cf. [Gär88]), instead of an expansion. Both points are made clear in the following example.

3.28. EXAMPLE. Consider the Kripke model $\mathcal{M} = < \mathcal{S}, \pi, \mathrm{R}, \mathbf{r}, \mathbf{c} >$ where

- $\mathcal{S} = \{s_0, s_1, s_2\}$,
- $\pi(p, s_j) = \mathbf{0} = \pi(q, s_i) \Leftrightarrow j = i = 2$,
- $\mathrm{R}(i) = \mathcal{S} \times \mathcal{S}$,

- **r** and **c** are arbitrary.

Now $\mathcal{M}, s_0 \models (\neg \mathbf{K}_i p \wedge \neg \mathbf{K}_i q)$ and $\mathcal{M}, s_0 \models <\text{do}_i(\text{test } p)> (\mathbf{K}_i p \wedge \mathbf{K}_i q)$, i.e, by performing a test for p knowledge concerning the formula q is acquired. Also $\mathcal{M}, s_0 \models \mathbf{K}_i \neg \mathbf{K}_i p$ and $\mathcal{M}, s_0 \models <\text{do}_i(\text{test } p)> \mathbf{K}_i \mathbf{K}_i p$, i.e, the test for p results in knowledge revision with regard to the formula $\neg \mathbf{K}_i p$.

Since the demand for minimal change is not formalized as informativeness and truthfulness are, no analogon of the first and the third case of lemma 3.25 exist for minimal change. However with regard to updates in a **KT** setting we can prove the following proposition.

3.29. PROPOSITION (Minimal change in a **KT** setting). *Let* $\mathcal{M} =< \mathcal{S}, \pi, \mathrm{R}, \mathbf{r}, \mathbf{c} >$ *be a* **KT** *model. Let* $s \in \mathcal{S}$, $\mathcal{M}' =< \mathcal{S}', \pi', \mathrm{R}', \mathbf{r}', \mathbf{c}' >$, *and* $\mathcal{M}', s' = \mathbf{r}(i, \text{test } \varphi)(\mathcal{M}, s)$. *Then:*

- $\mathcal{S}' = \mathcal{S}$, $\pi' = \pi$, $\mathbf{r}' = \mathbf{r}$, $\mathbf{c}' = \mathbf{c}$, *and* $s' = s$,
- $\mathrm{R}'(j) = \mathrm{R}(j)$ *for* $j \neq i$,
- *Let* $\mathcal{M}'' =< \mathcal{S}'', \pi'', \mathrm{R}'', \mathbf{r}'', \mathbf{c}'' >$ *be a* **KT** *model such that* $\mathcal{M}'' \in \mathcal{M}_{\subseteq}$. *Then it holds that*

$$\mathcal{M}'', s \models (\mathbf{K}_i \varphi \vee \mathbf{K}_i \neg \varphi) \Rightarrow \mathrm{R}''(i) \subseteq \mathrm{R}'(i)$$

PROOF OF PROPOSITION 3.29: Let $\mathcal{M}' =< \mathcal{S}', \pi', \mathrm{R}', \mathbf{r}', \mathbf{c}' >$ be such that $(\mathcal{M}', s) = \mathbf{r}(i, \text{test } \varphi)(\mathcal{M}, s)$. From definition 3.6, it follows directly that $\mathcal{S}' = \mathcal{S}, \pi' = \pi, \mathbf{r}' = \mathbf{r}, \mathbf{c}' = \mathbf{c}$, $s' = s$, and $\mathrm{R}'(j) = \mathrm{R}(j)$ for $j \neq i$. Hence the first two clauses are straightforward. This leaves the third clause to be proved. To prove this clause we start with proving the following lemma.

3.30. LEMMA. *Let* $\mathcal{M}'$, $\mathcal{M}$, *and* s *be as above, and let* $\varphi \in \mathcal{L}_0$. *Let furthermore* $\mathcal{M}''$ *be in* $\mathcal{M}_{\subseteq}$. *It holds that:*

- $\forall t \in \mathcal{S}[\mathcal{M}, t \models \varphi \Leftrightarrow \mathcal{M}'', t \models \varphi]$.
- $\mathcal{M}, s \models \varphi \Rightarrow (\{t' \in \mathcal{S}' \mid (s, t') \in \mathrm{R}'(i)\} = \{t \in \mathcal{S} \mid (s, t) \in \mathrm{R}(i) \& \mathcal{M}, t \models \varphi\})$.
- $\forall s', t \in \mathcal{S}[((s', t) \in \mathrm{R}(i) \& s' \neq s) \Rightarrow (s', t) \in \mathrm{R}'(i)\}$.

PROOF OF LEMMA 3.30:

- By definition of epistemic submodels, it holds that $\pi' = \pi$ and $\mathcal{S}' = \mathcal{S}$. The first clause of the lemma now follows by induction on the formulae of $\mathcal{L}_0$.
- Assume that $\mathcal{M}, s \models \varphi$. By definition, $\mathrm{R}'(i) = \mathrm{R}(i) \backslash \{(s, s') \mid \text{not}(\mathcal{M}, s \models \varphi \Leftrightarrow \mathcal{M}, s' \models \varphi)\}$. From $\mathcal{M}, s \models \varphi$, it then follows that $\mathrm{R}'(i) = \mathrm{R}(i) \backslash \{(s, s') \mid \mathcal{M}, s' \not\models \varphi\}$. Thus $\{t' \in \mathcal{S}' \mid (s, t) \in \mathrm{R}'(i)\} = \{t \in \mathcal{S} \mid (s, t) \in \mathrm{R}(i) \& \mathcal{M}, t \models \varphi\}$.
- By definition, $\mathrm{R}'(i) = \mathrm{R}(i) \backslash \{(s, s') \mid \text{not}(\mathcal{M}, s \models \varphi \Leftrightarrow \mathcal{M}, s' \models \varphi)\}$, i.e, the only pairs (u, v) that are removed from $\mathrm{R}(i)$ are those with $u = s$. Thus if $(s', t) \in \mathrm{R}(i)$ and $s' \neq s$, then also $(s', t) \in \mathrm{R}'(i)$.

⊠

Since $\mathcal{M}'', s \models \mathbf{K}_i\varphi \vee \mathbf{K}_i\neg\varphi$, it follows by definition of $\models$ that $\mathcal{M}'', s \models \mathbf{K}_i\varphi$ or $\mathcal{M}'', s \models \mathbf{K}_i\neg\varphi$. We show that if $\mathcal{M}'', s \models \mathbf{K}_i\varphi$ it follows that $\mathrm{R}''(i) \subseteq \mathrm{R}'(i)$; the case where $\mathcal{M}'', s \models \mathbf{K}_i\neg\varphi$ is then completely analogous.

Hence assume that $\mathcal{M}'', s \models \mathbf{K}_i\varphi$. To show: for all $(u, v) \in \mathrm{R}''(i)$, it holds that $(u, v) \in \mathrm{R}'(i)$. We distinguish two cases:

$u \neq s$: Now since $\mathrm{R}''(i) \subseteq \mathrm{R}(i)$, it follows by the third clause of lemma 3.30 that $(u, v) \in \mathrm{R}'(i)$.

$u = s$: From $\mathcal{M}'', s \models \mathbf{K}_i\varphi$, it follows that $\mathcal{M}'', v \models \varphi$. Since $\varphi \in \mathcal{L}_0$, it follows by the first clause of lemma 3.30 that $\mathcal{M}, v \models \varphi$. Since $\mathcal{M}''$ is a $\mathbf{KT}$ model, $\mathrm{R}''(i)$ is reflexive and hence $\mathcal{M}'', s \models \varphi$. Again applying the first clause of the lemma yields that $\mathcal{M}, s \models \varphi$. Now from $\mathcal{M}, s \models \varphi, \mathcal{M}, v \models \varphi$ and $\mathrm{R}''(i) \subseteq \mathrm{R}(i)$, it follows by the second clause of lemma 3.30 that $(s, v) = (u, v) \in \mathrm{R}'(i)$.

The proposition hence holds in case $\mathcal{M}, s \models \mathbf{K}_i\varphi$; as stated above, the case where $\mathcal{M}, s \models \mathbf{K}_i\neg\varphi$ is completely analogous.

⊠

Proposition 3.29 states that our semantics for `test` in a **KT** setting is such that these tests cause the *globally minimal change*. With this it is meant that the model that results from performing a test is in every aspect as similar as possible to the model in which the test is performed. Intuitively this is clear: if informativeness is to be ensured at least the epistemic alternatives that prevent informativeness need to be removed. Since exactly those alternatives are removed in our approach, and nothing else changes, we get an absolutely minimal change.

For tests in a **S5** setting, proposition 3.29 does not hold. Example 3.31 shows why this is the case.

3.31. EXAMPLE. Consider the **S5** model $\mathcal{M}$ that contains four states, two of which satisfy p, the other two satisfy $\neg p$. The `r` and `c` functions are arbitrary. The leftmost part of figure 3.2 contains a graphical version of the valuation and the non-reflexive part of the epistemic accessibility relation for some agent i. Now assume that agent i performs a test for p in the state s. If we implement tests as we did it in definition 3.15, the second model as given in figure 8 results, and 8 pairs (u, v) are removed from the set $\mathrm{R}(i)$. However a different model $\mathcal{M}''$ with $\mathcal{M}'', s \models \mathbf{K}_i p$ exists, namely the rightmost model in figure 8, for which only 6 pairs (u, v) need to be removed.

Two important remarks have to be made concerning example 3.31. Firstly, note that the alternative implementation of epistemic updates as exemplified by the rightmost model, is a rather counterintuitive one. As remarked in section 2, tests for a certain proposition φ should result in knowledge concerning the truth of φ. The approach that we take towards this acquisition of knowledge is that the information that the agent conceives by performing a test results in the rejection of the epistemic alternatives of the agent that contradict this new information. This is, mutatis mutandis, exactly what happens in the second of the models of example 3.31. However, in the third model *all* epistemic alternatives are rejected, whether they support the new information or not. This seems
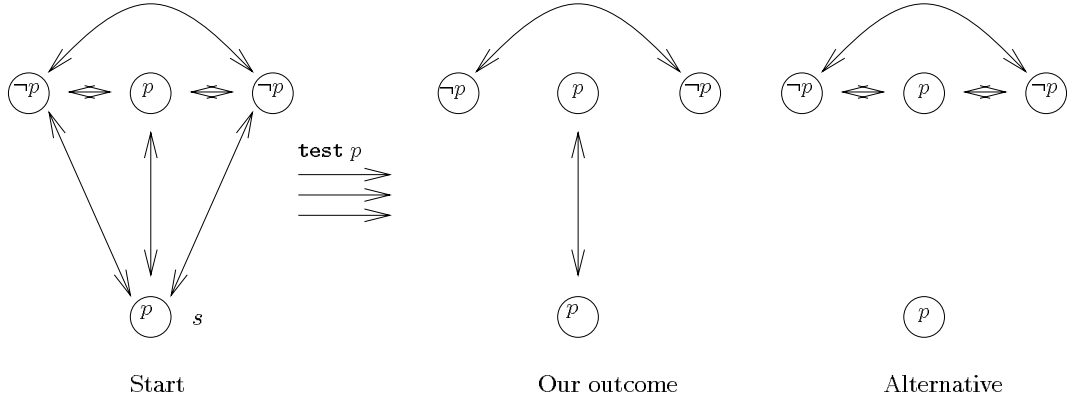
19

Figure 8: Different implementations of **S5** tests.

to be a rather questionable move: the incoming information is not such that all epistemic alternatives can be declared incorrect, but only those that contradict this new information. From an intuitive point of view, the alternative implementation suggested in example 3.31 is unacceptable.

Another important remark following example 3.31, is given by the fact that the alternative approach results in a model, counterintuitive though it might be, in which less pairs are removed and such that the demand for informativeness is met. But this seems to question the minimal change that our approach was supposed to ensure. However we claim that with regard to the changes that really matter, our approach indeed ensures minimal change. For when looked upon from the point of view of the agent that performs the test, the changes caused by the test in our **S5** approach are indeed minimal: the only change that this agent may observe is the extension of his/her knowledge, and this extension is minimal given the demand for informativeness. In other words, the tests in the **S5** setting as we implemented them can be said to cause *locally minimal* changes, this in contrast with the *globally minimal* changes that **KT** tests cause. Intuitively, global minimality should imply local minimality (the **KT** tests indeed cause locally minimal change), and locally minimal change seems to be the really relevant notion.

3.32. PROPOSITION (Minimal change in an **S5** setting). *Let* $\mathcal{M} =< \mathcal{S}, \pi, \mathrm{R}, \mathbf{r}, \mathbf{c} >$ *be an* **S5** *model. Let* $s \in \mathcal{S}$, $\mathcal{M}' =< \mathcal{S}', \pi', \mathrm{R}', \mathbf{r}', \mathbf{c}' >$, *and* $\mathcal{M}', s' = \mathbf{r}(i, \texttt{test } \varphi)(\mathcal{M}, s)$. *Then:*

- $\mathcal{S}' = \mathcal{S}$, $\pi' = \pi$, $\mathbf{r}' = \mathbf{r}$, $\mathbf{c}' = \mathbf{c}$, *and* $s = s'$,
- $\mathrm{R}'(j) = \mathrm{R}(j)$ *for* $j \neq i$,
- *Let* $\mathcal{M}'' =< \mathcal{S}'', \pi'', \mathrm{R}'', \mathbf{r}'', \mathbf{c}'' >$ *be an* **S5** *model such that* $\mathcal{M}'' \in \mathcal{M}_{\subseteq}$, *and* $\mathcal{M}'', s \models (\mathbf{K}_i \varphi \vee \mathbf{K}_i \neg \varphi)$. *Then it holds that* $[s]_{\mathrm{R}''(i)} \subseteq [s]_{\mathrm{R}'(i)}$.

PROOF OF PROPOSITION 3.32: Let $\mathcal{M}' =< \mathcal{S}', \pi', \mathrm{R}', \mathbf{r}', \mathbf{c}' >$ be such that $(\mathcal{M}', s) = \mathbf{r}(i, \texttt{test } \varphi)(\mathcal{M}, s)$. From definition 3.15, it follows directly that $\mathcal{S}' = \mathcal{S}, \pi' = \pi, \mathbf{r}' = \mathbf{r}, \mathbf{c}' = \mathbf{c}, s' = s$, and $\mathrm{R}'(j) = \mathrm{R}(j)$ for $j \neq i$. Hence the first two clauses are straightforward.

20

Since $\mathcal{M}'', s \models \mathbf{K}_i\varphi \vee \mathbf{K}_i\neg\varphi$, it holds that $\mathcal{M}'', s \models \mathbf{K}_i\varphi$ or $\mathcal{M}'', s \models \mathbf{K}_i\neg\varphi$. We show the case where $\mathcal{M}'', s \models \mathbf{K}_i\varphi$; the other case is completely analogous.

We start by proving the first two clauses of lemma 3.30 for an **S5** setting.

3.33. LEMMA. *Let $\mathcal{M}'$, $\mathcal{M}$, and s be as above, and let $\varphi \in \mathcal{L}_0$. Let furthermore $\mathcal{M}''$ be in $\mathcal{M}_{\subseteq}$. It holds that:*

- $\forall t \in \mathcal{S}[\mathcal{M}, t \models \varphi \Leftrightarrow \mathcal{M}'', t \models \varphi]$.
- $\mathcal{M}, s \models \varphi \Rightarrow (\{t' \in \mathcal{S}' \mid (s,t) \in \mathrm{R}'(i)\} = \{t \in \mathcal{S} \mid (s,t) \in \mathrm{R}(i) \,\&\, \mathcal{M}, t \models \varphi\})$.

PROOF OF LEMMA 3.33:   The first clause is proved as in the proof of lemma 3.30. For the second clause note that $\mathrm{R}'(i)$ is defined as $(\mathrm{R}(i) \setminus \mathrm{Cl}_{\mathrm{eq}}([s]_{\mathrm{R}(i)})) \cup \mathrm{Cl}_{\mathrm{eq}}([s]_{\mathrm{R}(i)}^{\varphi+}) \cup \mathrm{Cl}_{\mathrm{eq}}([s]_{\mathrm{R}(i)}^{\varphi-})$. From this definition, it follows that $(s,t) \in \mathrm{R}'(i)$ iff $(s,t) \in \mathrm{Cl}_{\mathrm{eq}}([s]_{\mathrm{R}(i)}^{\varphi+})$, for all $t \in \mathcal{S}$. Now assume that $\mathcal{M}, s \models \varphi$. Then:

$$
\begin{aligned}
& (s,t) \in \mathrm{R}'(i) \\
\Leftrightarrow\ & (s,t) \in \mathrm{Cl}_{\mathrm{eq}}([s]_{\mathrm{R}(i)}^{\varphi+}) \\
\Leftrightarrow\ & (s,t) \in \mathrm{R}(i) \text{ and } (\mathcal{M}, s \models \varphi \Leftrightarrow \mathcal{M}, t \models \varphi) \\
\Leftrightarrow\ & (s,t) \in \mathrm{R}(i) \text{ and } \mathcal{M}, t \models \varphi
\end{aligned}
$$

This suffice to conclude that the lemma holds.
⊠

Now let $t \in \mathcal{S}''$ be such that $(s,t) \in \mathrm{R}''(i)$. Since $\mathrm{R}''(i) \subseteq \mathrm{R}(i)$, also $(s,t) \in \mathrm{R}(i)$. Since $\mathcal{M}'', s \models \mathbf{K}_i\varphi$, it holds that $\mathcal{M}'', t \models \varphi$. Now since $\varphi \in \mathcal{L}_0$, it follows by the first clause of lemma 3.33 that also $\mathcal{M}, t \models \varphi$, and $\mathcal{M}, s \models \varphi$. Since $(s,t) \in \mathrm{R}(i)$, it follows by the second clause of lemma 3.33 that $(s,t) \in \mathrm{R}'(i)$, which suffices to prove the case for $\mathcal{M}'', s \models \mathbf{K}_i\varphi$.
⊠

Proposition 3.32 states that the changes implied on the model are minimal when looked upon from the state in which the test is performed: for this state the number of epistemic alternatives is maximal given the fact that knowledge is acquired. This is exactly as desired: the world surrounding the agent in the state in which the test is performed is identical to the world before the test was performed in as many aspects as possible. The only notion that is essentially affected by the test is the knowledge of the agent, and locally minimal change ensures that the change in knowledge is as little as possible.

Besides the general demands of informativeness, truthfulness, and minimal change, our test actions also satisfy the additional postulates given in section 2.1. This fact is expressed using a variant of the action transformations as introduced in [HLM93a].

3.34. DEFINITION. Two actions α and α' are a-equivalent, notation $\alpha \equiv_a \alpha'$, iff for all $\mathcal{M}, s$ it holds that

- $\lambda i.\mathrm{r}(i, \alpha)(\mathcal{M}, s) = \lambda i.\mathrm{r}(i, \alpha')(\mathcal{M}, s)$.
- $\lambda i.\mathrm{c}(i, \alpha)(\mathcal{M}, s) = \lambda i.\mathrm{c}(i, \alpha')(\mathcal{M}, s)$.

3.35. LEMMA. *For all formulae $\varphi \in \mathcal{L}_0$ and for all agents i we have:*

21

- if $\mathbf{K}_i\varphi \vee \mathbf{K}_i\neg\varphi$ then test φ else skip fi $\equiv_a$ skip.
- if $\mathbf{K}_i\varphi \vee \mathbf{K}_i\neg\varphi$ then skip else test φ fi $\equiv_a$ test φ.
- test $\varphi \equiv_a$ test φ; test φ.
- test $\varphi \equiv_a$ test $\neg\varphi$.

The first two clauses of lemma 3.35 formalize the idea that tests in a state of a priori knowledge come down to empty actions. The third clause states that tests are idempotent: a sequence of two tests for φ behaves essentially the same as a single test for φ. The last clause of lemma 3.35 corresponds to the idea that testing for the truth of a proposition is equivalent to testing for the falsity.

The following lemma formalizes some observations made in example 3.11.

3.36. LEMMA. *Let φ and ψ be formulae from $\mathcal{L}_0$, let i be some agent, and let $\mathcal{M}$ be some Kripke model.*

- test $\varphi \equiv_a$ test ψ *if φ and ψ are equivalent in propositional logic.*
- $\mathcal{M}, s \models \varphi \wedge \psi \Rightarrow \mathbf{r}(i, \text{test } \varphi \wedge \psi)(\mathcal{M}, s) = \mathbf{r}(i, \text{test } \varphi; \text{test } \psi)(\mathcal{M}, s)$.

Summarizing, the semantics that we proposed in section 3 is such that our test actions are not only tests in the general sense, but also in the more restricted AI sense: all the demands given in section 2 (that is informativeness, truthfulness, and minimal change) and in 2.1 (a priori knowledge, idempotence, extension of concrete knowledge, neutrality of phrasing) are met by the tests as we formalized them.

Besides the properties of tests as given above, an interesting relation exists with the notion of a selection function as used in Stalnaker's approach to conditional logic ([Sta68]). Given a Kripke model $\mathcal{M}$, the selection function $f(\varphi, s)$ gives the *most preferred* or *most reasonable* world given φ and s. Stalnaker gives four restrictions that a reasonable selection function should meet:

(1) φ is true at $f(\varphi, s)$.
(2) $f(\varphi, s)$ is undefined only if s' is inaccessible from s for all worlds s' in which φ holds.
(3) If φ is true at s then $f(\varphi, s) = s$.
(4) If φ is true at $f(\psi, s)$ and ψ is true at $f(\varphi, s)$ then $f(\varphi, s) = f(\psi, s)$.

It turns out that, given some obvious modifications, these restrictions are met by our definition of $\mathbf{r}(i, \text{test } \varphi)(\mathcal{M}, s)$, i.e., by testing a transition takes place to the most reasonable world in which $\mathbf{K}_i\varphi \vee \mathbf{K}_i\neg\varphi$ holds. This result depends on locally minimal change only, and holds therefore both in a **KT** and an **S5** setting.

3.37. DEFINITION. Let $\mathcal{M} = < \mathcal{S}, \pi, \mathrm{R}, \mathbf{r}, \mathbf{c} >$ be some Kripke model.

- The function $\mathrm{kn} : \mathcal{A} \times \mathcal{L}_0 \to \mathcal{L}$ is defined by:

$$\mathrm{kn}(i, \varphi) = \mathbf{K}_i\varphi \vee \mathbf{K}_i\neg\varphi$$

- The (partial) function $f_k : \mathcal{L} \times \mathcal{M}_{\subseteq} \times \mathcal{S} \to \mathcal{M}_{\subseteq} \times \mathcal{S}$ is defined by:

$$f_k(\mathrm{kn}(i, \varphi), \mathcal{M}', s) = \mathbf{r}(i, \text{test } \varphi)(\mathcal{M}', s)$$

3.38. LEMMA. *Let $\mathcal{M} \, = \, < \, \mathcal{S}, \pi, \mathrm{R}, \mathrm{r}, \mathrm{c} \, >$ be some Kripke model. For all $i \in \mathcal{A}$, $\varphi \in \mathcal{L}_0$, $\mathcal{M}' \in \mathcal{M}_{\subseteq}$ and $s \in \mathcal{S}$ it holds:*

(1) $\mathrm{kn}(i, \varphi)$ *is true at* $f_k(\mathrm{kn}(i, \varphi), \mathcal{M}', s)$.

(2) $f_k(\mathrm{kn}(i, \varphi), \mathcal{M}', s)$ *is defined.*

(3) *If* $\mathrm{kn}(i, \varphi)$ *is true at* $\mathcal{M}', s$ *then* $f_k(\mathrm{kn}(i, \varphi), \mathcal{M}', s) = \mathcal{M}', s$.

(4) *If* $\mathrm{kn}(i, \varphi)$ *is true at* $f_k(\mathrm{kn}(i, \psi), \mathcal{M}', s)$ *and* $\mathrm{kn}(i, \psi)$ *is true at* $f_k(\mathrm{kn}(i, \varphi), \mathcal{M}', s)$ *then* $f_k(\mathrm{kn}(i, \psi), \mathcal{M}', s) = f_k(\mathrm{kn}(i, \varphi), \mathcal{M}', s)$.

PROOF OF LEMMA 3.38: The first three clauses are consequences of the theorems and lemmas given above. In particular, clause 1 corresponds to informativeness of our tests, clause 2 is due to the fact that tests are serial (both clauses are consequences of the second item of lemma 3.25), and clause 3 is a case of a priori knowledge (first two items of lemma 3.35). We show the fourth clause. Hence assume that $\mathrm{kn}(i, \varphi)$ is true at $f_k(\mathrm{kn}(i, \psi), \mathcal{M}', s)$ and $\mathrm{kn}(i, \psi)$ is true at $f_k(\mathrm{kn}(i, \varphi), \mathcal{M}', s)$. We now distinguish four cases, namely $\mathcal{M}', s \models \varphi \wedge \psi$, $\mathcal{M}', s \models \varphi \wedge \neg\psi$, $\mathcal{M}', s \models \neg\varphi \wedge \psi$ and $\mathcal{M}', s \models \neg\varphi \wedge \neg\psi$, and show the first case; the other three cases are analogous. If $\mathcal{M}', s \models \varphi \wedge \psi$, then from the fact that $\mathrm{kn}(i, \varphi)$ is true at $f_k(\mathrm{kn}(i, \psi), \mathcal{M}', s)$ it follows from the fact that tests are truthful that $\mathrm{r}(i, \mathtt{test}\ \psi)(\mathcal{M}', s) \models \mathbf{K}_i\varphi \wedge \mathbf{K}_i\psi$. This latter fact implies that $\mathcal{M}', s \models \mathbf{K}_i(\psi \to \varphi)$. Analogously it follows from the fact that $\mathrm{kn}(i, \psi)$ is true at $f_k(\mathrm{kn}(i, \varphi), \mathcal{M}', s)$, that $\mathcal{M}', s \models \mathbf{K}_i(\varphi \to \psi)$, and thus $\mathcal{M}', s \models \mathbf{K}_i(\varphi \leftrightarrow \psi)$. Hence the epistemic alternatives of s in $\mathcal{M}'$ that satisfy φ are exactly those epistemic alternatives that satisfy ψ. It then follows from theorems 3.29 and 3.32 that the models that are given by $f_k(\mathrm{kn}(i, \psi), \mathcal{M}', s)$ and $f_k(\mathrm{kn}(i, \varphi), \mathcal{M}', s)$ are identical.
⊠

4. Tests as belief expansions

In [Gär88], Peter Gärdenfors defined some postulates for belief expansion, belief revision, and belief contraction respectively. Belief expansions are considered as modelling the epistemic change that follows learning something. As Gärdenfors states it:

> 'The most common causes of such [epistemic] changes are observations and the information provided by other people.' ([Gär88],p. 48)

Given the fact that our sort of tests are meant to formalize observations, it seems rather natural to investigate the relation between tests and the Gärdenfors postulates for belief expansion.

When formalizing belief expansions, Gärdenfors starts from the notion of a *belief set*, this is a consistent, deductively closed set of sentences that models the beliefs of an agent. A belief set is typically denoted by K, and the expansion of K with some formula φ is denoted by K^+_φ. With regard to belief expansions, Gärdenfors defines the following.

4.1. DEFINITION. The Gärdenfors postulates for belief expansion:

(K^+1) K^+_φ is a belief set.

$(\mathrm{K}^+\,2)$ $\varphi \in \mathrm{K}^+_\varphi$.

$(\mathrm{K}^+\,3)$ $\mathrm{K} \subseteq \mathrm{K}^+_\varphi$.

$(\mathrm{K}^+\,4)$ if $\varphi \in \mathrm{K}$ then $\mathrm{K}^+_\varphi = \mathrm{K}$.

$(\mathrm{K}^+\,5)$ if $\mathrm{K} \subseteq \mathrm{H}$ then $\mathrm{K}^+_\varphi \subseteq \mathrm{H}^+_\varphi$.

$(\mathrm{K}^+\,6)$ For all belief sets K, and all sentences φ, K^+_φ is the smallest set that satisfies $(\mathrm{K}^+\,1)$ – $(\mathrm{K}^+\,5)$.

If we take K^+_φ to be related to the knowledge of an agent after a test is performed, and K to be related to the knowledge before execution of the test, then postulate 2 can intuitively be seen as corresponding to *informativeness*, postulate 3 corresponds to expansion of the knowledge of the agent concerning propositional formulae, postulate 4 is analogous to *idempotence*, and postulates 5 and 6 can be seen as having a certain 'minimal change' flavour.

In our formal system, the *knowledge sets*, of agents are defined with regard to a given state in a given model.

4.2. DEFINITION. Let $\mathcal{M}$ be some model, let s be some state in $\mathcal{M}$, and let i be some agent.

- The knowledge set of the agent i in $\mathcal{M}, s$, notation $\mathrm{K}(i, \mathcal{M}, s)$, is defined by

$$\mathrm{K}(i, \mathcal{M}, s) = \{\psi \in \mathcal{L}_0 \mid \mathcal{M}, s \models \mathbf{K}_i \psi\}$$

- The expansion of $\mathrm{K}(i, \mathcal{M}, s)$ with the formula $\varphi \in \mathcal{L}_0$, notation $\mathrm{K}^+_\varphi(i, \mathcal{M}, s)$, is defined by

$$\mathrm{K}^+_\varphi(i, \mathcal{M}, s) = \{\psi \in \mathcal{L}_0 \mid \mathcal{M}, s \models [\mathrm{do}_i(\mathtt{test}\ \varphi)]\mathbf{K}_i \psi\}$$

4.3. LEMMA. *Let $\mathcal{M}$ be some Kripke model, let s be some state in $\mathcal{M}$, and let $\varphi \in \mathcal{L}_0$ be some formula. Then $\mathrm{K}^+_\varphi(i, \mathcal{M}, s) = \mathrm{K}(i, \mathcal{M}', s)$, where $\mathcal{M}', s = \mathtt{r}(i, \mathtt{test}\ \varphi)(\mathcal{M}, s)$.*

An important difference between expansions with ψ as Gärdenfors sees them, and tests for ψ as we see them, is given by the 'neutrality' of our sort of tests, i.e., a test for φ has the same effect as a test for $\neg\varphi$, this in correspondence with the common sense intuition concerning tests. Taking into account his neutrality, the first four Gärdenfors postulates can be proved both for the **KT** and **S5** implementations of our tests. The function pol, for polarity, is defined to let the notational similarity between our versions of the Gärdenfors postulates and the original ones, be as large as possible.

4.4. DEFINITION. Let $\mathcal{M}$ be some Kripke model. The function pol is defined by:

$$
\begin{aligned}
\mathrm{pol} \quad &: \ (\mathcal{M}_{\subseteq} \times \mathcal{S}) \to \mathcal{L}_0 \to \mathcal{L}_0 \\
\mathrm{pol}(\mathcal{M}', s)(\varphi) &= \varphi \text{ if } \mathcal{M}', s \models \varphi \\
&= \neg\varphi \text{ if } \mathcal{M}'s, \models \neg\varphi
\end{aligned}
$$

Usually, $\mathrm{pol}(\mathcal{M}, s)(\varphi)$ is denoted by $\mathrm{pol}_{\mathcal{M},s}(\varphi)$.

4.5. THEOREM. *Let $\mathcal{M}$ be either a* **KT** *or an* **S5** *model, let s be some state in $\mathcal{M}$, and let i be some agent. The following is valid for all $\varphi \in \mathcal{L}_0$.*

(K^+1) $\mathrm{K}_\varphi^+(i, \mathcal{M}, s)$ *is a belief set.*

(K^+2) $\mathrm{pol}_{\mathcal{M},s}(\varphi) \in \mathrm{K}_\varphi^+(i, \mathcal{M}, s)$.

(K^+3) $\mathrm{K}(i, \mathcal{M}, s) \subseteq \mathrm{K}_\varphi^+(i, \mathcal{M}, s)$.

(K^+4) *if* $\mathrm{pol}_{\mathcal{M},s}(\varphi) \in \mathrm{K}(i, \mathcal{M}, s)$ *then* $\mathrm{K}_\varphi^+(i, \mathcal{M}, s) = \mathrm{K}(i, \mathcal{M}, s)$.

PROOF OF THEOREM 4.5:

(K^+1) Due to the seriality of the epistemic accessibility relation, it is straightforward to prove that for all models $\mathcal{M}$ and states s, $\mathrm{K}(i, \mathcal{M}, s)$ is consistent and deductively closed. Since $\mathbf{r}(i, \mathtt{test}\ \varphi)(\mathcal{M}, s)$ is also serial (cf. case 2 of lemma 3.25), some pair $(\mathcal{M}', s)$ exists such that $\mathcal{M}', s = \mathbf{r}(i, \mathtt{test}\ \varphi)(\mathcal{M}, s)$. Then $\mathrm{K}_\varphi^+(i, \mathcal{M}, s) = \mathrm{K}(i, \mathcal{M}', s)$, and hence $\mathrm{K}_\varphi^+(i, \mathcal{M}, s)$ indeed is a belief set.

(K^+2) Since our tests are informative, it holds that $\mathcal{M}, s \models < \mathrm{do}_i(\mathtt{test}\ \varphi) > \mathbf{K}_i \varphi$ or $\mathcal{M}, s \models < \mathrm{do}_i(\mathtt{test}\ \varphi) > \mathbf{K}_i \neg \varphi$. Since the epistemic accessibility relation is reflexive, it follows that either $\varphi \in \mathrm{K}_\varphi^+(i, \mathcal{M}, s)$ or $\neg \varphi \in \mathrm{K}_\varphi^+(i, \mathcal{M}, s)$.

(K^+3) By the fourth case of lemma 3.25, tests increase the knowledge of the agent concerning aspects of the world. Hence for all $\psi \in \mathcal{L}_0$, if $\psi \in \mathrm{K}(i, \mathcal{M}, s)$ then $\psi \in \mathrm{K}_\varphi^+(i, \mathcal{M}, s)$, and hence $\mathrm{K}(i, \mathcal{M}, s) \subseteq \mathrm{K}_\varphi^+(i, \mathcal{M}, s)$.

(K^+4) In lemma 3.35 it is stated that tests satisfy the first of the postulates given in definition 2.1. Hence if $\mathcal{M}, s \models \mathbf{K}_i \varphi$ or $\mathcal{M}, s \models \mathbf{K}_i \neg \varphi$, then $\mathbf{r}(i, \mathtt{test}\ \varphi)(\mathcal{M}, s) = \mathcal{M}, s$, which suffices to conclude that (K^+4) holds.

⊠

In order to prove the last two postulates for the epistemic updates caused by our tests, we have to make some rather weak assumptions regarding the models under consideration. We demand that the two knowledge sets that are compared correspond to models that have a common parent, i.e, a model that is an epistemic 'super model' of both these models. This assumption needs to be made since our definition of knowledge set is a model dependent one, this in contrast with the definition of belief sets as given in [Gär88].

4.6. THEOREM. *Let $\mathcal{M}$ be either a* **KT** *or an* **S5** *model, let $\mathcal{M}', \mathcal{M}'' \in \mathcal{M}_\subseteq$, let s be some state in $\mathcal{M}$, and let i be some agent. The following is valid for all $\varphi \in \mathcal{L}_0$.*

(K^+5) *If* $\mathrm{K}(i, \mathcal{M}', s) \subseteq \mathrm{K}(i, \mathcal{M}'', s)$, *then* $\mathrm{K}_\varphi^+(i, \mathcal{M}', s) \subseteq \mathrm{K}_\varphi^+(i, \mathcal{M}'', s)$.

(K^+6) $\mathrm{K}_\varphi^+(i, \mathcal{M}, s)$ *is the smallest set that satisfies the postulates 1 – 4 as given in theorem 4.5 and the postulate 5 as given in this theorem. More in particular, if $\varphi \in \mathrm{K}(i, \mathcal{M}'', s)$ or $\neg \varphi \in \mathrm{K}(i, \mathcal{M}'', s)$, then $\mathrm{K}_\varphi^+(i, \mathcal{M}, s) \subseteq \mathrm{K}(i, \mathcal{M}'', s)$.*

PROOF OF THEOREM 4.6:

(K^+5) We prove this case by contraposition. Assume that $\mathrm{K}_\varphi^+(i, \mathcal{M}', s) \not\subseteq \mathrm{K}_\varphi^+(i, \mathcal{M}'', s)$. Let $\psi \in \mathcal{L}_0$ be such that $\psi \in \mathrm{K}_\varphi^+(i, \mathcal{M}', s)$ and $\psi \notin \mathrm{K}_\varphi^+(i, \mathcal{M}'', s)$. Let $\mathcal{M}_a'$[4] be such that $\mathbf{r}(i, \mathtt{test}\ \varphi)(\mathcal{M}', s) = \mathcal{M}_a', s$, and let $\mathcal{M}_a''$ be such that $\mathbf{r}(i, \mathtt{test}\ \varphi)(\mathcal{M}'', s) = \mathcal{M}_a'', s$. We distinguish two cases:

---

[4]The a stands for 'after'.

(1) $\mathcal{M}', s \models \varphi$. Then $\mathcal{M}'_a, s \models \mathbf{K}_i\varphi$. Since $\psi \in \mathrm{K}^+_\varphi(i, \mathcal{M}', s)$, it follows that $\mathcal{M}'_a, t \models \psi$, for all t with $(s, t) \in \mathrm{R}'_a(i)$. Since $\psi \in \mathcal{L}_0$, it follows by the first clause of lemmas 3.30 and 3.33 that $\mathcal{M}', t \models \psi$, for all t with $(s, t) \in \mathrm{R}'_a(i)$. By clause 2 of lemmas 3.30 and 3.33 we know that $\{s' \mid (s, s') \in \mathrm{R}'_a(i)\} = \{s' \mid (s, s') \in \mathrm{R}'(i) \,\&\, \mathcal{M}', s' \models \varphi\}$. Thus $\mathcal{M}', t \models \psi$ for all t such that $(s, t) \in \mathrm{R}'(i) \,\&\, \mathcal{M}', t \models \varphi$. But this implies that $\mathcal{M}', t \models (\varphi \to \psi)$, for all t with $(s, t) \in \mathrm{R}'(i)$. Hence $\mathcal{M}', s \models \mathbf{K}_i(\varphi \to \psi)$, and thus $(\varphi \to \psi) \in \mathrm{K}(i, \mathcal{M}', s)$.

Since $\psi \notin \mathrm{K}^+_\varphi(i, \mathcal{M}'', s)$, $\mathcal{M}'', s \not\models \mathbf{K}_i\psi$, and it follows that some t exists with $(s, t) \in \mathrm{R}''_a(i)$ and $\mathcal{M}''_a, t \models (\neg\psi \wedge \varphi)$. By the first clause of lemmas 3.30 and 3.33, it follows that $\mathcal{M}'', t \models (\neg\psi \wedge \varphi)$. Since $\mathcal{M}''_a \in \mathcal{M}''_\subseteq$, it holds that $(s, t) \in \mathrm{R}''(i)$. Thus $\mathcal{M}'', t \models (\neg\psi \wedge \varphi)$ for some t with $(s, t) \in \mathrm{R}''(i)$. Hence $\mathcal{M}'', s \not\models \mathbf{K}_i(\varphi \to \psi)$, and thus $(\varphi \to \psi) \notin \mathrm{K}(i, \mathcal{M}'', s)$. Since $(\varphi \to \psi) \in \mathrm{K}(i, \mathcal{M}', s)$ and $(\varphi \to \psi) \notin \mathrm{K}(i, \mathcal{M}'', s)$, it follows that $\mathrm{K}(i, \mathcal{M}', s) \not\subseteq \mathrm{K}(i, \mathcal{M}'', s)$, which was to be shown.

(2) $\mathcal{M}', s \models \neg\varphi$. This case is completely analogous to the case where $\mathcal{M}', s \models \varphi$, replacing φ by $\neg\varphi$.

This suffices to conclude $(\mathrm{K}^+ 5)$.

$(\mathrm{K}^+ 6)$ From propositions 3.29 and 3.32 it follows that the set of epistemic alternatives from s in the model $\mathcal{M}'$, which is the result of performing a test for φ in $(\mathcal{M}, s)$, is maximal given the fact that $\mathbf{K}_i\varphi$ or $\mathbf{K}_i\neg\varphi$ holds in $\mathcal{M}', s$. Hence if some model $\mathcal{M}'' \in \mathcal{M}_\subseteq$ satisfies $\mathcal{M}'', s \models \mathbf{K}_i\varphi \vee \mathbf{K}_i\neg\varphi$ then the set of epistemic alternatives of s in $\mathcal{M}''$ is a subset of the set of epistemic alternatives of s in $\mathcal{M}'$. Thus the set $\mathrm{K}(i, \mathcal{M}'', s)$ is a superset of $\mathrm{K}(i, \mathcal{M}', s)$.

$\boxtimes$

In [Gär88] some additional properties of belief expansions are given, that presuppose that the language under consideration allows conjunction and implication of sentences.

4.7. LEMMA ([Gär88]). *The following are consequences of the basic postulates given in definition 4.1.*

(1) *If* $\vdash \varphi \leftrightarrow \psi$, *then* $\mathrm{K}^+_\varphi = \mathrm{K}^+_\psi$.
(2) $(\mathrm{K}^+_\varphi)^+_\psi = \mathrm{K}^+_{\varphi \wedge \psi}$.
(3) $(\mathrm{K}^+_\varphi)^+_\psi = (\mathrm{K}^+_\psi)^+_\varphi$.

In order to prove these additional properties for the knowledge expansions caused by our tests actions, some care has to be taken concerning the truth or falsity of the conjunction in relation with truth or falsity of its conjuncts. First an additional definition is given that is used to mimic successive expansions.

4.8. DEFINITION. Let $\mathcal{M}$ be either a **KT** or an **S5** model, let s be some state in $\mathcal{M}$, and let i be some agent. Then $(\mathrm{K}^+_{\varphi_1})^+_{\varphi_2}(i, \mathcal{M}, s)$ is defined by

$$(\mathrm{K}^+_{\varphi_1})^+_{\varphi_2}(i, \mathcal{M}, s) = \{\psi \in \mathcal{L}_0 \mid \mathcal{M}, s \models [\mathrm{do}_i(\texttt{test } \varphi_1; \texttt{test } \varphi_2)]\mathbf{K}_i\psi\}$$

4.9. THEOREM. *Let* $\mathcal{M}$ *be either a* **KT** *or an* **S5** *model, let* s *be some state in* $\mathcal{M}$, *and let* i *be some agent. Let* $\varphi_1, \varphi_2 \in \mathcal{L}_0$, *and let* $\vdash_{\mathrm{pl}}$ *denote derivability in classical propositional logic.*

(1) *If $\vdash_{\mathrm{pl}} \varphi_1 \leftrightarrow \varphi_2$ then $\mathrm{K}^+_{\varphi_1}(i, \mathcal{M}, s) = \mathrm{K}^+_{\varphi_2}(i, \mathcal{M}, s)$.*

(2) $\mathrm{K}^+_{\varphi_1 \wedge \varphi_2}(i, \mathcal{M}, s)$ *is in general* <u>*not*</u> *equal to* $(\mathrm{K}^+_{\varphi_1})^+_{\varphi_2}(i, \mathcal{M}, s)$.

(3) $\mathrm{K}^+_{\mathrm{pol}_{\mathcal{M},s}(\varphi_1) \wedge \mathrm{pol}_{\mathcal{M},s}(\varphi_2)}(i, \mathcal{M}, s) = (\mathrm{K}^+_{\varphi_1})^+_{\varphi_2}(i, \mathcal{M}, s)$.

(4) $(\mathrm{K}^+_{\varphi_1})^+_{\varphi_2}(i, \mathcal{M}, s) = (\mathrm{K}^+_{\varphi_2})^+_{\varphi_1}(i, \mathcal{M}, s)$.

PROOF OF THEOREM 4.9: Case 1 is straightforward and left to the reader. Case 2 follows from the observations made in example 3.11. Cases 3 and 4 are proved using the following lemma.

4.10. LEMMA. *Let $\mathcal{M} = < \mathcal{S}, \pi, \mathrm{R}, \mathbf{r}, \mathbf{c} >$ be some Kripke model, $s \in \mathcal{S}$, $i \in \mathcal{A}$, and let $\varphi, \psi \in \mathcal{L}_0$. Let $\mathcal{M}^1, s = \mathbf{r}(i, \mathtt{test}\ \varphi; \mathtt{test}\ \psi)(\mathcal{M}, s)$, $\mathcal{M}^2, s = \mathbf{r}(i, \mathtt{test}\ \psi; \mathtt{test}\ \varphi)(\mathcal{M}, s)$, and $\mathcal{M}^3, s = \mathbf{r}(i, \mathtt{test}\ \mathrm{pol}_{\mathcal{M},s}(\varphi) \wedge \mathrm{pol}_{\mathcal{M},s}(\psi))(\mathcal{M}, s)$. Then:*

$$(s, t) \in \mathrm{R}^1(i) \Leftrightarrow (s, t) \in \mathrm{R}^2(i) \Leftrightarrow (s, t) \in \mathrm{R}^3(i)$$

PROOF OF LEMMA 4.10: We distinguish four cases: $\mathcal{M}, s \models \varphi \wedge \psi, \mathcal{M}, s \models \varphi \wedge \neg\psi, \mathcal{M}, s \models \neg\varphi \wedge \psi$, and $\mathcal{M}, s \models \neg\varphi \wedge \neg\psi$. We show the first case; all other cases are completely analogous. Hence assume that $\mathcal{M}, s \models \varphi \wedge \psi$. Let $\mathcal{M}', s = \mathbf{r}(i, \mathtt{test}\ \varphi)(\mathcal{M}, s)$. By lemmas 3.30 and 3.33, it follows that $(s, t) \in \mathrm{R}'(i)$ iff $(s, t) \in \mathrm{R}(i)\ \&\ \mathcal{M}, t \models \varphi$, and $(s, t) \in \mathrm{R}^1(i)$ iff $(s, t) \in \mathrm{R}'(i)\ \&\ \mathcal{M}', t \models \psi$. Then:

$$\begin{aligned}
&(s, t) \in \mathrm{R}^1(i) \\
\Leftrightarrow\ &(s, t) \in \mathrm{R}'(i) \text{ and } \mathcal{M}', t \models \psi && \text{by case 2 of 3.30 and 3.33} \\
\Leftrightarrow\ &(s, t) \in \mathrm{R}'(i) \text{ and } \mathcal{M}, t \models \psi && \text{by case 1 of 3.30 and 3.33} \\
\Leftrightarrow\ &(s, t) \in \mathrm{R}(i) \text{ and } \mathcal{M}, t \models \varphi \text{ and } \mathcal{M}, t \models \psi \\
\Leftrightarrow\ &(s, t) \in \mathrm{R}(i) \text{ and } \mathcal{M}, t \models \mathrm{pol}_{\mathcal{M},s}(\varphi) \text{ and } \mathcal{M}, t \models \mathrm{pol}_{\mathcal{M},s}(\psi) && \text{since } \mathcal{M}, s \models \varphi \wedge \psi \\
\Leftrightarrow\ &(s, t) \in \mathrm{R}(i) \text{ and } \mathcal{M}, t \models \mathrm{pol}_{\mathcal{M},s}(\varphi) \wedge \mathrm{pol}_{\mathcal{M},s}(\psi) \\
\Leftrightarrow\ &(s, t) \in \mathrm{R}^3(i) && \text{by case 2 of 3.30 and 3.33}
\end{aligned}$$

In a completely analogous way $(s, t) \in \mathrm{R}^2(i) \Leftrightarrow (s, t) \in \mathrm{R}^3(i)$ is proved, which suffices to conclude lemma 4.10.
⊠

Now let $\mathcal{M}_a, s = \mathbf{r}(i, \mathtt{test}\ \varphi_1; \mathtt{test}\ \varphi_2)(\mathcal{M}, s)$, i.e., $(\mathrm{K}^+_{\varphi_1})^+_{\varphi_2}(i, \mathcal{M}, s) = \mathrm{K}(i, \mathcal{M}_a, s)$, and let $\mathcal{M}'_a = \mathbf{r}(i, \mathtt{test}\ \mathrm{pol}_{\mathcal{M},s}(\varphi_1) \wedge \mathrm{pol}_{\mathcal{M},s}(\varphi_2))(\mathcal{M}, s)$. From lemma 4.10 it follows that $(s, t) \in \mathrm{R}_a(i) \Leftrightarrow (s, t) \in \mathrm{R}'_a(i)$, for all $(s, t) \in \mathrm{R}(i)$. Hence $\mathrm{K}(i, \mathcal{M}_a, s) = \mathrm{K}(i, \mathcal{M}'_a, s)$, which suffices to conclude case 3 of theorem 4.9.

Let $\mathcal{M}''_a, s = \mathbf{r}(i, \mathtt{test}\ \varphi_2; \mathtt{test}\ \varphi_1)(\mathcal{M}, s)$. Then from lemma 4.10 it follows that $(s, t) \in \mathrm{R}_a(i) \Leftrightarrow (s, t) \in \mathrm{R}''_a(i)$, for all $(s, t) \in \mathrm{R}(i)$, which suffices to conclude case 4 of theorem 4.9.
⊠

4.11. REMARK. Note that the second and the third clause of theorem 4.9 are completely in accordance with the observations made in example 3.11. In terms of example 3.11, the second clause corresponds to the idea that if a pretty and dumb agent asks the magical mirror whether s/he is smart and pretty, this only results in the agent knowing that s/he is either not smart or not pretty. Should s/he have asked successively whether s/he is

smart and whether s/he is pretty, s/he would know as a result that s/he is pretty but not smart. The third clause corresponds to a smart and pretty agent asking whether s/he is smart and pretty: this leads to the knowledge that s/he is indeed smart and pretty; the same knowledge would have resulted if s/he had asked successively whether s/he is smart and whether s/he is pretty.

The strong correspondence between belief expansions à la Gärdenfors and our sort of tests, gives us another means to investigate the minimization of the change caused by performing a test. Analogously to [Gär88], we can prove that the knowledge set of the agent is minimally expanded, given the fact that it is deductively closed and it contains knowledge concerning the proposition that is tested for.

4.12. LEMMA. *Let $\mathcal{M} = < \mathcal{S}, \pi, \mathrm{R}, \mathbf{r}, \mathbf{c} >$ be some Kripke model, let $s \in \mathcal{S}$, and let $\varphi \in \mathcal{L}_0$. Then:*

$$\mathrm{K}_\varphi^+(i, \mathcal{M}, s) = \mathrm{Cn}(\mathrm{K}(i, \mathcal{M}, s) \cup \{\mathrm{pol}_{\mathcal{M},s}(\varphi)\})$$

where Cn denotes deductive closure in classical propositional logic.

PROOF OF LEMMA 4.12: We prove that the two sets are equal by proving that each set is a subset of the other.

'$\supseteq$' By $(\mathrm{K}^+ 3)$ we have that $\mathrm{K}(i, \mathcal{M}, s) \subseteq \mathrm{K}_\varphi^+(i, \mathcal{M}, s)$, and by $(\mathrm{K}^+ 2)$, $\{\mathrm{pol}_{\mathcal{M},s}(\varphi)\} \subseteq \mathrm{K}_\varphi^+(i, \mathcal{M}, s)$. From $(\mathrm{K}^+ 1)$ it follows that $\mathrm{K}_\varphi^+(i, \mathcal{M}, s)$ is deductively closed, and hence $\mathrm{Cn}(\mathrm{K}(i, \mathcal{M}, s) \cup \{\mathrm{pol}_{\mathcal{M},s}(\varphi)\}) \subseteq \mathrm{K}_\varphi^+(i, \mathcal{M}, s)$.

'$\subseteq$' Note that if $\psi \in \mathrm{K}_\varphi^+(i, \mathcal{M}, s)$, then $\mathrm{pol}_{\mathcal{M},s}(\varphi) \to \psi \in \mathrm{K}(i, \mathcal{M}, s)$, by an analogous argument as given in the proof of theorem 4.6. But then $\{\mathrm{pol}_{\mathcal{M},s}(\varphi) \to \psi, \mathrm{pol}_{\mathcal{M},s}(\varphi)\} \subseteq \mathrm{Cn}(\mathrm{K}(i, \mathcal{M}, s) \cup \{\mathrm{pol}_{\mathcal{M},s}(\varphi)\})$, and hence $\psi \in \mathrm{Cn}(\mathrm{K}(i, \mathcal{M}, s) \cup \{\mathrm{pol}_{\mathcal{M},s}(\varphi)\})$.

⊠

From the theorems and lemmas given in this section, it follows that both the **KT** and the **S5** implementations of our tests, which were directed at satisfying the basic postulates and the additional AI postulates as given in section 2, furthermore satisfy (slightly modified versions of) the Gärdenfors postulates for belief expansions. Hence one might conclude that our sorts of tests formalize observations as they are looked upon by Gärdenfors, which provides an additional argument for the statement that our tests behave in an intuitively acceptable and desirable way.

5. Related work

In this section we look at two fairly well-known examples of formalizations of 'tests' as these occur in the literature. In particular we will try and determine whether these 'tests' actually are tests in the sense of section 2.

5.1. Tests in dynamic logic

In dynamic logic tests are well-known actions that are usually considered to be among the set of standard actions (cf. [Har84, KT90, Gol92]). The dynamic logic tests, or dl-test for short, check for the truth of a given formula in a given state: if the formula is true, nothing happens, else the test action fails and the program aborts. Although this approach is perfectly usable when the end state behaviour of programs is considered, and although some nice technical results follow, this definition of dl-tests fails to formalize the intuitive meaning of tests as we gave it in section 2.

Firstly, dl-tests, when considered in an epistemic/dynamic framework as given in [HLM93a, HLM93b], are in general not informative: if the dl-test for φ succeeds, then the resulting state is the begin state, and hence if the agent did not have any information concerning φ, s/he will not have any information after the test is performed. If the dl-test fails, no end state results, and hence the dl-test would be vacuously informative according to the formalization given in definition 3.18.

Dl-tests are truthful: if φ holds then dl-testing for φ would not change anything, and in particular φ would still hold. If φ does not hold (and hence $\neg\varphi$ holds), then testing for φ would lead to an empty set of end states, and hence $\neg\varphi$ is vacuously true in all end states. Thus according to definition 3.21 this would mean that dl-tests are indeed truthful. However, the fact that falsity of φ would lead to an empty set of end states when testing for φ, is the reason why these dl-test are not strongly truthful.

With regard to minimal change, we find that in case of failure dl-tests do most certainly not meet this demand: if the dl-test fails, no end state results and therefore one cannot say that the dl-test causes minimal change to the begin state. All in all, the dl-tests are by no means tests as we define them in section 2.

It is argued by several authors that given the discrepancy between the test actions from dynamic logic and human, common sense tests, the former ones should be renamed. In [HLM93a], as well as in [HLM93b] and in this paper, we use the name *confirmations* to refer to the tests from dynamic logic; Krister Segerberg ([Seg93]) suggested to use *verify* to refer to dynamic logic tests, since in his opinion verification is what actually takes place.

5.2. Robert Moore's tests

In [Moo84] Robert Moore defined a framework for knowledge and actions, one of the main goals of which was to give an intuitively acceptable formalization of common sense human testing. Moore's idea of a test is the following.

> 'The essence of a test is that it is an action with a directly observable result that depends conditionally on an unobservable precondition.' ([Moo84], p. 2).

It is obvious that this is a different starting point than we have: we treat tests as observations, and deal with an eventual unobservability through the abilities of agents. Moore gives three properties that an action should have for it to be called a test:

(1) after an agent performs the action s/he should know whether the observable result holds.

(2) after the agent performs the action s/he should know that s/he had just performed the action[5].

(3) the agent knows that the observable result holds after the test is performed just in case that the observable precondition was true before the test was performed.

These demands differ substantially from the demands that we formulated in section 2. The demands for truthfulness and minimal change are not implied by Moore's properties. Informativeness holds to the extent that in some successor state knowledge is acquired concerning the state in which the test was performed. However due to the fact that neither truthfulness nor minimal change is demanded it needs not be the case that this knowledge is still of use in the successor state.

In the concrete example that is given in [Moo84], viz. that of a test for the acidity of a solution using blue litmus paper, a demand for truthfulness is made in order to ensure that the knowledge of the agent still applies. The test that is used in this example is a very specific one (a litmus test) and not an instance of a general one like the tests in our formalization.

Summarizing, due to the different emphasis that is placed in [Moo84], the tests considered there differ substantially from tests as we see them. In particular Moore's notion of tests cannot be used to formalize the idea of tests as observations as we introduced it in this paper.

6. Discussion

In this paper we looked into the nature of tests as a means to acquire knowledge. We gave three demands that actions should meet for them to be called tests: informativeness, truthfulness, and the demand for minimal change. Furthermore some additional postulates are defined that are characteristic for tests performed in typical AI environments. A possible formalization of a general test action `test` φ is given. This formalization incorporates some elements of update semantics into the framework for events and abilities that we defined in [HLM93a, HLM93b]. Two implementations of tests are given, one in a **KT** and one in an **S5** setting. For both implementations it holds that the `test` action satisfies both the general demands for tests and the additional demands imposed on AI tests. Furthermore, both the implementation of tests as we give them are such that the Gärdenfors postulates for belief expansion are satisfied, which supports the idea that our kind of tests models the observations made by agents in an intuitively correct way.

Our further research on the topics introduced in this paper will be focussed on the treatment of defeasibility (cf. for instance [Bes89] or [Łuk90]): if knowledge, or better belief, is no longer infallible it should be possible to formalize the process of agents giving up certain believes. A possible implementation of this belief revision could be based on the idea of adding epistemic alternatives, analogously to the implementation of belief expansion by the removal of epistemic alternatives as we did it in this paper.

---

[5]It is not at all clear that it is actually possible to formalize this property in Moore's system, let alone to define actions that have this property.

Acknowledgements

Thanks are due to Frans Voorbraak for providing us with example 3.12.

References

[Bes89]   P. Besnard. *An Introduction to Default Logic.* Springer-Verlag, Berlin-Heidelberg, 1989.

[Che80]   B.F. Chellas. *Modal Logic. An Introduction.* Cambridge University Press, Cambridge, 1980.

[Dav90]   E. Davis. *Representation of Commonsense Knowledge.* Morgan Kaufmann, San Mateo, 1990.

[EV93]   J. van Eijck and F.-J. de Vries. Reasoning about update logic. Technical Report CS-R9312, CWI, Amsterdam, 1993. To appear in the Journal of Philosophical Logic.

[Gär88]   P. Gärdenfors. *Knowledge in Flux: Modeling the Dynamics of Epistemic States.* The MIT Press, Cambridge, Massachusetts and London, England, 1988.

[Gol92]   R. Goldblatt. *Logics of Time and Computation*, volume 7 of *CSLI Lecture Notes*. CSLI, Stanford, 1992. Second edition.

[Har84]   D. Harel. Dynamic logic. In D.M. Gabbay and F. Guenthner, editors, *Handbook of Philosophical Logic*, volume 2, chapter 10, pages 497–604. D. Reidel, Dordrecht, 1984.

[HLM93a]   W. van der Hoek, B. van Linder, and J.-J. Ch. Meyer. A logic of capabilities. Technical Report IR-330, Vrije Universiteit Amsterdam, July 1993. An extended abstract appeared in [HS93].

[HLM93b]   W. van der Hoek, B. van Linder, and J.-J. Ch. Meyer. Unravelling nondeterminism: On having the ability to choose. Technical Report RUU-CS-93-30, Utrecht University, September 1993.

[HM85]   J.Y. Halpern and Y.O. Moses. A guide to the modal logics of knowledge and belief. In *Proc. 9th IJCAI*, pages 480–490, 1985.

[HR83]   J. Halpern and J. Reif. The propositional dynamic logic of deterministic, well-structured programs. *Theoretical Computer Science*, 27:127–165, 1983.

[HS93]   J. Horty and Y. Shoham, Program Chairs. Reasoning about mental states: Formal theories & applications. Technical Report SS-93-05, AAAI Press, 1993. Papers from the 1993 Spring Symposium.

[KL88]   S. Kraus and D. Lehmann. Knowledge, belief and time. *Theoretical Computer Science*, 58:155–174, 1988.

31

[KT90]   D. Kozen and J. Tiuryn. Logics of programs. In J. van Leeuwen, editor, *Handbook of Theoretical Computer Science*, volume B, pages 789–840. Elsevier, 1990.

[Łuk90]   W. Łukaszewicz. *Non-Monotonic Reasoning*. Series in Artificial Intelligence. Ellis Horwood, 1990.

[MH]   J.-J. Ch. Meyer and W. van der Hoek. Epistemic logic for AI and computer science. Manuscript.

[Moo84]   R.C. Moore. A formal theory of knowledge and action. Technical Report 320, SRI International, 1984.

[Seg93]   K. Segerberg. Lectures given at the fifth European summer school in Logic, Language, and Information. Lisbon, Portugal, August 1993.

[Sta68]   R. Stalnaker. A theory of conditionals. In N. Rescher, editor, *Studies in Logical Theory*, number 2 in American Philosophical Quarterly Monograph Series, pages 98–122. Blackwell, Oxford, 1968.

[Vel91]   F. Veltman. Defaults in update semantics. Technical Report LP-91-02, ILLC, University of Amsterdam, 1991.