

# Language for Description of Worlds. Part 1: Theoretical Foundation

Dimitar Dobrev

Institute of Mathematics and Informatics,  
Bulgarian Academy of Sciences, Bulgaria  
[d@dobrev.com](mailto:d@dobrev.com)

## Abstract

We will reduce the task of creating AI to the task of finding an appropriate language for description of the world. This will not be a programming language because programming languages describe only computable functions, while our language will describe a somewhat broader class of functions. Another specificity of this language will be that the description will consist of separate modules. This will enable us look for the description of the world automatically such that we discover it module after module. Our approach to the creation of this new language will be to start with a particular world and write the description of that particular world. The point is that the language which can describe this particular world will be appropriate for describing any world. This is the first part of the paper. In this part we will define the basic theoretical concepts which will we will need when creating the language we are looking for.

*Keywords:* Artificial General Intelligence, Language for description of worlds, Event-Driven Model, Definition of Algorithm

*ACM Computing Classification System 2012:* Computing methodologies → Artificial intelligence → Philosophical/theoretical foundations of artificial intelligence

*Mathematics Subject Classification 2020:* 68T01

## 1 Introduction

This paper presents a new approach to the exploration of AI. This is the Event-Driven (ED) approach. The underlying idea of the ED approach is that instead of absorbing all input/output information, the model should reflect only the events which matter (“important events”).

Every action is an event. Every observation is an event, too. If the model were to reflect each and every action and observation, it would end up overloaded with an enormous amount of information. The overloaded situation can however be avoided when the model is limited only to certain important events. This leads us to the idea of Event-Driven models.

A disadvantage (or perhaps an advantage) of the ED model is that it does not describe the world completely, but only partially. More precisely, the ED model describes a certain class of worlds (the worlds which comply with a certain pattern).

**What makes this paper different?** The mainstream approach to dealing with multi-agent systems is based on the assumption that the world is given (known) and what we need to find is a policy. In other words, the world is part of the known terms of the problem while the policy is the unknown part. This paper is different because we will assume that instead of being given, the world is unknown and has to be found.

The assumption that the world is given implies that we have a relation which provides a full description of the world. Conversely, in this paper we will try to provide partial descriptions of the world and will do so by employing ED models.

**Structure of the language:** The description of the world will not be similar to a homogenous system consisting of one single layer. Our description of the world would rather be structured as a multilayer system. Figure 1 presents our multilayer structure as a pyramid where the first layer is the base of the pyramid.

Thus, events will form the first layer of our system. We will use these events in order to describe the Event-Driven (ED) models. Importantly, the states of the ED models must not be the same (some states must be different). For this condition to be fulfilled, there must be at least one state with a special occurrence in it so that the state can be distinguished from other states. We will name this special occurrence *distinction*. The set of the distinctions we will designate as *trace*.

Our first idea about the trace is that it is fixed. (For example, let us have a state in which it is always cold.) Then our notion of a trace will evolve and we

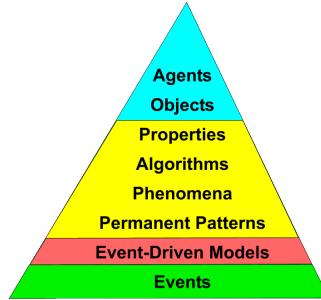


Figure 1: Structure of the language

will have a *moving* trace which can appear and then disappear or move from one state to another. (In our example the chill could move to the room next door.)

The next layers of the pyramid will consist of various patterns (we will present all of them by ED models). We will begin with the permanent patterns, i.e. those which are observed all the time. While most authors assume that all patterns are permanent, in this paper we reckon that in addition to the permanent ones there are other, impermanent patterns which are observed only from time to time. We will name these impermanent patterns *phenomena*. In other words, just as traces can be *fixed* or *moving*, patterns can also be *permanent* or *phenomena*.

Algorithms are also impermanent patterns (observed only when an algorithm is being executed). We will present algorithms as sequences of events. Typically, algorithms are understood as sequences of *actions* on the basis of the assumption that the protagonist is always the one who executes the algorithm. In this paper the algorithm will be any pattern and if the actions of an agent are aimed at maintaining the pattern, then we will say that the agent is the one who executes the algorithm.

When a phenomenon is associated with the observation of an object, we will call that phenomenon a *property*. This takes us to an abstraction of a higher order, namely the abstractive concept of *object*. Objects are not directly observable, but are still identifiable through their properties.

The next abstraction we get to will be *agents*. Similar to objects, we cannot observe agents directly, but can still gauge them on the basis of their actions.

In order to describe the world, we have to describe the agents which live that world and explain what we know about these agents. The most important thing to describe about agents is whether they are our friends or foes and accordingly what will they try to do to us by their actions – help us or disrupt us?

The descriptions above relate to computable worlds (ones that can be emulated by a computer program). While any presence of a non-computable agent in the world would make the world itself non-computable, there is another way to make non-computable worlds. We may add a rule which depends on the existence of some algorithm (more precisely, on the existence of an execution of that algorithm). The question “Does an execution of the algorithm exist?” is non-computable (halting problem [1]).

### Contributions

1. Event-Driven model. (The concept has already been introduced by [2], but that paper did not provide an interpretation of the ED model. This is important, because interpretations are what makes models meaningful and distinguish between adequate and inadequate models.)

2. This paper proves that Markov Decision Process (MDP) is a special case of an ED model and that an ED model is the natural generalization of MDP.

3. Simple MDP. We will simplify the MDP to obtain a more straightforward model which can describe more worlds.

4. Extended model. This is the model in which the state knows everything. We will use this model in order to introduce an interpretation of events and ED models. (Although the Extended model was introduced in [3], in that version of the model the state knows only what has happened and what is going to happen, but does not know what is possible to happen. Therefore, the state of the Extended model in [3] knows nothing about the missed opportunities. In [3] the Extended model is referred to as “maximal”.)

**How is this paper organized.** Sections 2 to 7 will provide the theoretical foundation of the paper. We will define the concepts *lived experience*, *life*, *Event-Driven model*, *phenomena*, *event* and *world*. We will give meaning to these concepts by defining *interpretation*. We will prove that MDP is a special case of an ED model.

## 2 Lived experience

The task is to describe the world on the basis of our lived experience. This means we are looking for a model which provides the best explanation of what has happened until the present moment. Hence the first question we will address

is “What is lived experience?”.

The common assumption is that we have only one life and lived experience comprises everything that has happened in that life until now. However, in this paper we will assume that we have more than one life and lived experience comprises everything that has happened both in the current life and in all previously lived lives. We will go even further and assume that we may have more than one current life.

Why do we make this assumption? From the individual’s perspective each individual has only one life, but from a population perspective there are many lives. In our case we have an artificial agent living in an artificial world. We can let the agent live many times in the world and will thus obtain many lives from which we can collect observations and try to identify patterns. We can even let multiple agents live in the world concurrently. Then will have multiple current lives.

**Definition 1.** Lived experience is a finite sequence of lives.

This takes us to the next question: “What is life?”

## 2.1 Life

*Life* will be a finite sequence of actions and observations. Life is what we see, but are there things in life which we do not see?

Let us imagine we have had two lives. In the first life we walked past a pot full of gold coins, but did not bother to open it and went on. In the second life we walked past an empty pot which we did not open either and again went on. In terms of actions and observations the two lives are fully identical, but they are still two different lives because in the first one we had the chance to find plenty of gold coins while in the second life we never had that chance.

That is why our description of life will consist of two determinants. The first one is what we have seen (the sequence of actions and observations). We will label this sequence as the *trace of life*.

**Note:** In this paper we use the terms *trace of live* and *trace of the ED model*. These are different things.

The second determinant of life is everything that was possible to happen in that life (regardless of whether it actually happened). The possible past and the possible future are described by the state of the world. Knowing what the state has been at each moment of time is important. That is why to the description of life we will add the sequence of states which the world has been through. We will label this sequence as the *backbone of life*.

The trace and the backbone describe life from the perspective of the world. If we look at the world from the agent's perspective, life will be described by the trace and by the guesses which the agent can make at any given moment. For example, at a given moment we can assume that after two steps we will see a certain observation. (The moment of the guess is the moment at which we are able to make that guess rather than the moment which the guess relates to).

Two things are important for each guess. First, the antecedents which must be true at the moment when we make the guess. The second one is the rule (heuristic) on the basis of which the guess can be made. We will assume that heuristics are derived from our entire lived experience (including the future lived experience). Hence, the addition of more lived experience leads to a change of heuristics and thereby to a change of guesses (even previous guesses). For example, there comes a moment when we realize that the sound produced by a fire weapon signifies some danger. Then we review our past and find out that every time we heard a fire shot we had been in trouble.

## 2.2 Heuristic

We said that in life there are things we do not see but which are still important enough for us to try predicting them. For example, "Are there gold coins in the pot?" We may try to predict this through our intuition or on the basis of some heuristics which we derive by collecting statistics from our lived experience. An example of such heuristic is "If the pot looks ancient, then it is full of gold coins". Certainly, each heuristic comes with some confidence coefficient. Let this coefficient be as low as 2%. Even with such a low confidence coefficient, if it appears to be an ancient pot it is worth the effort to open it and check for gold coins.

An heuristic will consist of antecedent, consequent and confidence coefficient. We will suppose that the consequent is an atomic formula, while the antecedent is a conjunction of atomic formulas. (We will define atomic formulas later.)

**Definition 2.** An heuristic is a material implication with confidence coefficient

$$A_1 \& A_2 \& \dots \& A_n \Rightarrow A_0 \quad (\textit{per cent}) \quad (1)$$

Here is an example of a heuristic:

$$Ob(t-1)=o_1 \& Act(t)=a \Rightarrow Ob(t+1)=o_2 \quad (10\%) \quad (2)$$

Here  $Ob(t) = o$  means that the observation at moment  $t$  is  $o$  and  $Act(t) = a$

means that the action at moment  $t$  is  $a$ . (We have expressed this as an equality because the action and the observation depend unambiguously on moment  $t$ .)

The event  $Ob(t)$  makes sense only at even-number moments, while  $Act(t)$  makes sense only at odd-number moments because we alternate actions and observations. The heuristic is valid for any  $t$ . That is why we will assume that the antecedent is known at moment  $t$  and this is the first moment in which the antecedent becomes known. (We can achieve this by replacing  $t$  with  $t + i$  for some  $i$ .)

The consequent here relates to a future visible event (predicting past visible events does not make sense). The consequent might also be related to an invisible events (past or future). For example:

$$Ob(t)=o_1 \Rightarrow PossNext(o_2, t-2) \quad (\text{some per cent}) \quad (3)$$

Here the invisible event  $PossNext(o, t)$  means it is possible that the next observation is  $o$ , i.e. it is possible that  $Ob(t + 2) = o$ . (In this expression we have not used equality because the possible next observations are more than one, while an equality implies one possibility.)

For  $o_1 = o_2$  the above heuristic will be meaningless because it will be tautologically true, while  $o_1 \neq o_2$  will make sense.

We will consider heuristics as rules which are always valid, but might also assume that a rule is valid only sometimes. In this case one impermanent rule will define one invisible event. That event will be True when the rule is valid and False when the rule is invalid. Let us consider the following rule:

$$\emptyset \Rightarrow Ob(t+2) \neq o \quad (4)$$

The antecedent here is the empty set, i.e. without antecedent or always. This rule tells us that the next observation cannot be  $o$ . This is sometimes True and sometimes False, but we do not see whether this is True (unless the next observation is  $o$ , but generally we cannot see that). Therefore, this is one invisible event and its negation is exactly  $PossNext(o, t)$ .

### 2.3 Atomic formula

The agent cannot observe all events and will have to stay with some finite number of events. The agent will assign a name to each observed event and we will call these events atomic formulas.

**Definition 3.** An atomic formula is an event to which we have assigned a name.

For example,  $Ob(t) = o$  is an atomic formula. (More precisely, these are multiple atomic formulas, one for each  $o$ .)

When observing some event we will actually be observing its negation as well. Thus, the negation of an atomic formula is also an atomic formula.

A conjunction of atomic formulas can also be an atomic formula if we have decided to observe it and provided that we have assigned a name to it. These conjunctions of course will be infinitely many and we would not be able to observe them all.

From event  $A(t)$  we can make countably many events  $A(t+i)$ , which will be the same event shifted in time. It does not make much sense to observe all these events because they are almost the same. Therefore we will observe only one of them. For example, if we have a conjunction of atomic formulas, we will decide to observe the event at the moment which is the maximum of the moments of the atomic formulas included in the conjunction.

In most cases the moment at which an observed event occurs is not important, but there are some exceptions. With ED models for example the exact time at which each of the observed events occurs is important (it is important whether an event occurs several steps earlier or several steps later).

## 2.4 Single-moment event

**Definition 4.** A single-moment event is one which depends only on one moment.

Let that moment be  $t$ . Then, a visible single-moment event is an event which depends only on one observation or only on one action (only on  $Ob(t)$  or only on  $Act(t)$ ). An invisible single-moment event is one which depends only on one state of the world ( $s_t$ ).

A conjunction of single-moment events can also be a single-moment event if all atoms in the conjunction are single-moment events and relate to one and the same moment.

The visible single-moment events are finitely many and there are as many invisible single-moment events as are the subsets of the states of the world (finite or continuum).

An atomic formula may or may not be a single-moment event.

## 2.5 The guess

At certain moments of time the agent will apply heuristics in order to derive some guesses. Each guess will consist of some consequent and some confidence



coefficient. The consequent will be an atomic formula and that formula will relate to a concrete moment of time  $t$  (which can be in the past or in the future). Hence, unlike the heuristic which applies to any  $t$ , the guess will apply to a particular  $t$ .

**Definition 5.** A guess is expressed as follows:

$$A(t) \quad (\textit{per cent}) \tag{5}$$

The set of guesses (heuristic applications) will be called *Guesses*.

Let us have another heuristic: “If the pot is full of gold coins, then we can buy an aircraft”. Let the confidence coefficient of this heuristic be 50%. Then as soon as we see a pot which appears ancient we apply the first heuristic and get “the pot is full of gold coins” (with a confidence coefficient of 2%). Now we can buy an aircraft (with a confidence coefficient of 1%).

The important takeaway from this example is that heuristics can be cascaded and that heuristic antecedents may include an invisible event. The cascade application of two heuristics can be considered as new one. However, it will take a lot of lived experience to find this new heuristic. The better way is to find more simple heuristics and combine them in more complex ones. We can find many heuristics which indicate that the pot is full of gold coins and their cascade application will lead to the consequent that in all these cases we can buy an aircraft. If we do not use cascade application then for every particular case we should collect individual statistics which will prove that in that particular case we can buy an aircraft.

## 2.6 The definition of life

In this definition life will be presented as a game between two players. The first player will be the agent (the protagonist), the second player will be “nature” and it will be the world itself. The agent called “nature” is not free to do whatever she likes. She operates to some rules, although she may have some freedom of choice (free will). Further down we will try to obtain an explanation of the world and will replace “nature” with a group of agents. That group will consist of zero, one, two or more agents. When dealing with the empty group we will reckon that in the world there are not any agents except the protagonist.

Life will consist of two things: trace and backbone.

**Definition 6.** Life is:

$$a_1, o_2, a_3, o_4, \dots, a_{k-1}, o_k, \quad \textit{where } a_i \in \Sigma, o_i \in \Omega. \tag{6}$$

$$u_0, w_1, u_2, w_3, \dots, w_{k-1}, u_k, \quad \text{where } u_i \in U, w_i \in W. \quad (7)$$

Here  $\Sigma$  are the possible actions,  $\Omega$  are the possible observations of the agent,  $U$  are the states of the world when it is the agent's turn to make a move and  $W$  are the states of the world when it is nature's turn to make a move. The symbol  $k$  stands for the length of life (if life is  $L$ , then its length is  $|L|$ ).

If we write the trace and the backbone in one line we get this:

$$u_0, a_1, w_1, o_2, u_2, a_3, w_3, \dots, a_{k-1}, w_{k-1}, o_k, u_k \quad (8)$$

From the perspective of the world life consists of a trace and a backbone, from the agent's perspective life consists of a trace and guesses and from their joint perspective life is life with guesses.

**Definition 7.** Life with guesses:

- trace
- backbone
- agent's guesses (heuristic applications):

$$g_0, g_1, g_2, \dots, g_k, \quad \text{where } g_i \subseteq \text{Guesses}. \quad (9)$$

The set  $g_i$  consists of the guesses which the agent can make at moment  $i$  (not earlier, but exactly at that moment). All valid guesses at moment  $t$  are:

$$G_t = \bigcup_{i \leq t} g_i \quad (10)$$

In some cases the set of guesses  $G_t$  may be contradictory. The guesses for some  $j$  can be both  $A(j)$  and  $\neg A(j)$  (perhaps with different confidence). The guesses  $A(j)$  and  $\neg A(j)$  may occur at the same moment (in one  $g_i$ ), but can also occur at two different moments.

The agent exercises thinking only when it is her turn to make a move, therefore only at even-number moments. If  $t$  is an even-number moment, the agent cannot choose the guesses  $g_t$  because they are defined by the past, but she can choose the guesses  $g_{t+1}$  because they are defined by the past plus the agent's next action. (In other words, the agent will run all possible actions and will choose the one which gives her the best  $g_{t+1}$ .)

### 3 Event-Driven models

Now we will sort of jump the gun and say what is an Event-Driven (ED) model before we even say what is an event *per se*. Roughly, an ED model is a directed graph where the nodes represent the states of the world and the arrows are labeled with events. We assume that the world keeps staying in the same state until some observed event occurs. Then the world shifts to another state by making a move over the arrow labeled with the event which has just occurred.

Distinctions are important elements of the ED model. A distinction in our language will mean that in a certain state a given event occurs with a probability different from the expected (average) probability. Distinctions will help us in two ways: i) understand in which state of the world we are in right now and ii) predict what is going to happen and what has happened. By *trace* we will mean the set of distinctions.

All arrows and distinctions in the ED model will have some probability. Instead of exact probabilities we will use probability intervals. First, we will explain why we prefer to use probability intervals. Second, we will explain why we make a distinction between things which are impossible and things which are possible but improbable. Third, we will introduce the statistics which will give meaning to these probabilities. And only then we will be able to provide the definition of ED models.

#### 3.1 The probability interval

In basketball, when you through the ball at the hoop, what is the probability of success? We may try 100 shots and calculate that probability by statistics. Other factors may also have some impact on that probability, such as the light conditions in the basketball hall or how fatigued you are. We may construct a more complex model which captures those factors as well, but some factors will always fall through the cracks and remain unaccounted for. These unaccounted factors will induce a degree of volatility meaning that the probability will not be an exact value but some interval.

The most important one of these additional factors is our freedom of choice (free will). When we shoot the ball, it matters a lot whether we want or do not want to score a goal. Let us assume that when we do not want the probability of scoring a goal by error is  $a$  and when we really want the probability of scoring (by intent) is  $b$ . Thus, each time we shoot at the hoop the probability of scoring a goal will be in the interval  $[a, b]$ .

### 3.2 Possible but improbable

Can there be some difference between the impossible and the improbable? We will make a distinction between a missing arrow and an arrow with a probability of zero (i.e. the interval  $[0, 0]$ ).

The probability of a given event may tend to zero. This for example is the case when an event occurs once in ten instances, then once in the next one hundred instances, etc. This kind of event is possible, but its probability is zero (or tends to zero).

When looking at the actions of the world (the observations of the agent) the difference between the impossible and the improbable is minor, but when looking at the agent's actions the difference becomes essential because this is the way to define the concept *incorrect move*.

When the arrows are labeled with actions and when an arrow is missing, then we will assume that the corresponding action is an incorrect move. This will be a separate event which we will designate as *semivisible*. An arrow with a probability of zero will mean that the action is correct but for some reason our agent will never (or almost never) execute that action.

### 3.3 Statistics

Statistics is the main tool which the agent uses in order to understand the world. Here we will deal with two types of statistics. The first one is the *statistics of the agent* and is based only on actions and observations. These are the real statistics available to the agent.

We will examine the second type of statistics. We will name them *statistics of the world*. In addition to actions and observations, these statistics capture the states of the world. We can imagine them as statistics created by some super agent who is capable of monitoring the state of the world. Certainly such a super agent does not exist. We assume that the world knows the state it is in. That is why the world will be able to create these statistics but will not be willing to share them with the agent.

Thus, the statistics of the world are something abstract which the agent cannot see. Nevertheless, we will use these statistics in order to define and give meaning to some of the introduced concepts. For example, what is the probability of moving from one state to another over a given arrow?

**Note:** We will use statistics of the world to determine the probabilities of arrows in the world and in the ED models. The problem is that we may have two different ED models that differ only in their probabilities. We assume

that the world has chosen one of these different ED models and determines its probabilities based on statistics.

Let us assume that the probabilities are fixed and we will calculate these fixed probabilities using the statistics of the world. We will make  $N$  steps and will count how many times we have been through each state and each arrow (for the case of ED models: the number of times we have been in the state and the event which labels the arrow has occurred). If we have been  $k$  times through a given arrow and  $m$  times through the tail (the state from which this arrow departs), then the probability of going through the arrow will be approximately  $k/m$ . Certainly, when the  $N$  value is low the error will be large, but when the  $N$  value is high we can assume that the probability is exactly  $k/m$ . Thus, the probability will be the limit value of  $k/m$  when  $N$  tends to infinity.

This is how we used the statistics of the world in order to give meaning to the probability of the arrows. Can we also use these statistics in order to define the probability intervals? This will be a bit more difficult. Let us again make  $N$  steps. Let  $n$  be the largest integer for which  $n^2 \leq N$ . We will break the interval  $N$  into  $n$  intervals such that in each interval there are at least  $n$  steps. We will derive statistics for each of these  $n$  intervals and will obtain some probability  $p_i$ . The probability interval of the arrow will be  $[\min(p_i), \max(p_i)]$ . When  $N$  tends to infinity we should obtain the exact value, but for this to be true we will make one assumption.

The assumption will be that the world and the agent change their fixed policies but do that only seldom, therefore any changes of the fixed policy of the ED model are also seldom (see below what is a fixed policy). When the world and the agent execute a fixed policy, then the probability of all arrows in the ED model will also be fixed. If the world and the agent change their policies too often, then this statistical method will not give us the result we need (i.e. will produce intervals which are shorter than the actual ones).

### 3.4 The ED model definition

**Definition 8.** An Event-Driven model consists of three parts:

Topology:

- $S$  – the set of the states of the model (finite or countable, i.e. most countable).
- $E$  – the set of the observed events (a limited number of events).

- $R \subseteq S \times E \times S$  – relation between the states, labeled with the observed events.
- Collision reconciliation rule.

Partial trace:

- $Trace \subseteq Distinctions$  – the set of distinctions.

Probabilities:

- Probability:  $R \rightarrow [0, 1] \times [0, 1]$  – a function which at each arrow returns a probability interval.
- Backward Probability: Same as Probability, but related to the past (when we move against the direction of the arrow).

**Note:** We have too many things in this definition. For each arrow we have some probability and even backward probability. In the cardinality of cases these probabilities will be absent. More precisely, the probability will be the interval  $[0, 1]$  which means that the probability is unknown.

**Definition 9.** Distinction is a 3-tuple consisting of state, event and probability. The probability is either a probability interval or the constant *never*. The probability must be different from the average probability of the event (the expected value). If the average probability of event  $e$  is  $[a_1, b_1]$  and the probability of the distinction is  $[a_2, b_2]$ , then  $[a_1, b_1] \not\subseteq [a_2, b_2]$ . (If  $[a_1, b_1] \subseteq [a_2, b_2]$ , then  $\langle s, e, [a_2, b_2] \rangle$  is not a distinction because, rather than being specified, the probability falls in an even broader range.)

We will assume that an arrow is missing then and only then when a distinction in the trace indicates that a given event cannot occur in a given state. Thus:

$$\forall s_1 \in S \quad \forall e \in E \quad \left( (\neg \exists s_2 \in S \quad \langle s_1, e, s_2 \rangle \in R) \Leftrightarrow \langle s_1, e, never \rangle \in Trace \right) \quad (11)$$

The observed events are those which sit over the arrows of the model. The events involved in the trace will be referred to as *additional events*. There may or may not be some intersection between the observed and the additional events.

A *trace* in our discussion is a *partial trace* because we assume that it contains only part of the distinctions of the model. If we assume that the trace holds all distinctions of the model, then it would be a full trace. (If the model has

more than one interpretation then we should be dealing with the trace of the interpretation rather than with the trace of the model.)

The collision reconciliation rule tells us which is the next state of the ED model in case that events  $a$  and  $b$  of  $E$  occur at the same time.

The rule can be deterministic, e.g.:

1. Event  $a$  has priority over event  $b$  and in this case the executed event will be  $a$ .
2. In this case the two events will be executed sequentially – first  $a$  and then  $b$ .
3. A separate arrow tells us where to go if  $a \& b$  occurs.

The rule can also be non-deterministic, e.g.:

4. We select  $a$  with a probability of  $p$  (or with a probability within the interval  $[p_1, p_2]$ ).

### 3.5 Interpretation

When is the ED model a model of the world? To answer this question, we will define the concepts *interpretation* and *characteristic*. For this definition we will use the statistics of the world. In other words, we will use the information about the state of the world – something which the agent does not know.

For the ED model to be a model of the world there must be a link between life and the state of the model. We will term this link an *interpretation* which, for any moment of life, will tell us (exactly or approximately) the state in which the ED model is. Hence, *interpretation* is a function which for each moment of life returns a state of the ED model (or a set of states or some *belief*).

An interpretation may not be unique, but if we expand the partial trace it will become unique. (We may expand the trace a little bit or expand it all the way to a full trace.)

Of course, interpretation is not an arbitrary function. It must be consistent with the ED model. When none of the observed events occurs in several consecutive moments, the interpretation function must return one and the same state for all these moments. If one of the observed events occurs, in the next moment the interpretation function must return a new state, provided that in the ED model there is an arrow over this event leading from the previous state to the new state.

We will consider several types of interpretations:

1. Simple interpretation. This is a function which for each state of the world returns a state of the ED model.

2. Unambiguous interpretation. This is a function which for each moment of an infinite life returns a state of the ED model. Each simple interpretation is unambiguous, but not vice versa.

3. Set-based interpretation. Same as the unambiguous interpretation, but returns a set of states of the ED model instead of a single set.

4. *Belief*-based interpretation. Same as the set-based interpretation, but returns some *belief* of states of the ED model (i.e. a set with probabilities).

In order to simplify the discussion, we will not deal with interpretations type 3 and 4. So, when we refer to an interpretation, we will mean an unambiguous interpretation.

The functions over the moments of life will be referred to as *multi-value events*. The usual events are two-value – they either occur or do not occur. If an event is  $n$ -value, we will present it as  $n$  two-value events which do not intersect (cannot occur at the same time). The interpretation is therefore a multi-value event. We will assume that the interpretation is defined for any infinite life. With finite lives we may have a problem defining some events. If we do not have information about the past or future of a certain life, the value of the event can be “I do not know”. For this reason, we assume that interpretations are defined over infinite lives.

**Definition 10.** An infinite life consists of an infinite trace and an infinite backbone (infinity is in both directions):

$$\dots, o_{-2}, a_{-1}, o_0, a_1, o_2, a_3, \dots \quad (12)$$

$$\dots, u_{-2}, w_{-1}, u_0, w_1, u_2, w_3, \dots \quad (13)$$

In one line it looks like this:

$$\dots, o_{-2}, u_{-2}, a_{-1}, w_{-1}, o_0, u_0, a_1, w_1, o_2, u_2, a_3, w_3, \dots \quad (14)$$

An interpretation described by visible events will be referred to as a visible interpretation. Each visible interpretation is unambiguous.

We assume that the ED model is described by its topology and its partial trace. On the basis of this description we look for an interpretation (one of the possible interpretations). Then, from the interpretation we can obtain the full trace and the probabilities.

### 3.6 Characteristic

Thus, an ED model must have an interpretation in order to be a model of the world, but this is not enough. That interpretation must have some meaningful characteristic.



Each event has its characteristic – this is a fuzzy set of the states of the world in which the event occurs.

**Definition 11.** The characteristic of an event is a function which for each state of the world gives the probability of that event to occur in that state.

The difference between *belief* and characteristic is that *belief* describes a single state while a characteristic describes set of states. The sum of the probabilities in *belief* is one, while in the case of a characteristic that sum is between zero and infinity.

An interpretation is a multi-value event, i.e. it consists of multiple events, each one with its own characteristic. The characteristic of an interpretation will be the set of all those characteristics. For the characteristic of an interpretation to be meaningful, all these characteristics must not be the same.

**Note:** We will assume that everything which the states of the world “remember” and “know” is important. (This is the assumption that the model of the world is minimal. Later on we will add this assumption.) Therefore, if the characteristics of two events are different there is something important which makes that difference.

**Note:** When the characteristic of an ED model is not meaningful, then the full trace is the empty set. That is, nothing interesting is going on.

The best case is when the characteristics are sets (rather than fuzzy sets). In this case we will have a simple interpretation and a meaningful characteristic.

It is important to know whether the model “remembers” important things. If it remembers only important things, then we will have a simple interpretation. If the model remembers only unimportant things, then we will not have an interpretation or will end up with an interpretation with a meaningless characteristic.

For example, let us take a Fully observable MDP (FOMDP). This ED model “remembers” the last observation. Let us assume that in our world this is something unimportant. The FOMDP interpretation is visible, but the characteristic of this unique interpretation will be meaningless because whatever the last observation is, the expected state of the world will be one and the same (due to our assumption that in our world this information is unimportant). If in our world the last observation carries some important information, then the characteristic of the FOMDP interpretation would be meaningful.

Let us now assume that the last observation is remembered by the state of the world (i.e. this is important information and each symbol in this information is important). Now that the world “remembers” the last observation, we can sort the states on the basis of this observation and obtain an equivalence relation. Then the FOMDP has a simple interpretation and it coincides with the visible

one.

**Note:** The above description applies to permanent patterns with a permanent trace. The pattern can also be transient (phenomenon) and in that case the interpretation will not apply to the full life, but only to certain parts of that life. (Phenomena do not occur all the time but only from time to time.) The moving trace also complicates the model.

### 3.7 Examples

Let us have event  $a$  and the ED model which remembers whether the occurrences of  $a$  add up to an even or odd number (Figure 2).

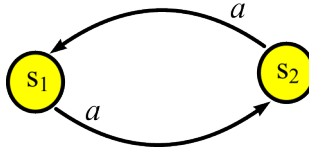


Figure 2: Count the event  $a$

Let us then assume that the odd or even number of occurrences of  $a$  is something important in our world. Then each state of the world “knows” this and we can sort the events in two sets (a set of even-number occurrences of  $a$  and all the rest). This gives us two simple interpretations of the ED model. We can relate the set of states in which the count of the occurrences of  $a$  is an even number to  $s_1$  or to  $s_2$ . Each of these two interpretations will give us a trace. We can take the full trace, but to distinguish the two interpretations it would be sufficient to take a random, non-empty part of the full trace. The arrow probabilities are also determined by the interpretation, but in this case all these probabilities are 1.

Now let us assume that the odd or even number of occurrences of  $a$  is something completely unimportant in our world. In this assumption the model in Figure 2 remembers only unimportant things. In this case the model will not have any interpretation because the interpretation needs something to start with. Neither does the state of the world know the state of model ( $s_1$  or  $s_2$ ), nor can we derive this information from the past and the future. Thus, there is not any function out there which can unambiguously tell between  $s_1$  and  $s_2$ .

If we waive the requirement that the model of the world remembers only important things, we may obtain a new model which will provide some inter-

pretation, however, that interpretation will tell us nothing (will not give us a trace). Let us expand the model of the world by adding that unimportant information (whether  $a$  occurred an even number of times). We will replace each state  $s$  with the states  $\langle s, \text{even} \rangle$  and  $\langle s, \text{odd} \rangle$ . If state  $s$  “remembers” that the count of  $a$  occurrences is an even number, then state  $\langle s, \text{even} \rangle$  will be reachable, while state  $\langle s, \text{odd} \rangle$  will be unreachable (as if it does not exist). If none of the states remembers that fact, then all these new states will be reachable. In this case  $s_1$  will relate to all states  $\langle s, \text{even} \rangle$  or to all states  $\langle s, \text{odd} \rangle$ . In both cases there will be no trace because the probability of each event will be the average probability.

For example, let us imagine our world in which we have the event “new day”. Sure the event is important, but whether it is an even-number day or an odd-number day is not important at all. What matters in this case is which day it is over a module of seven because this is the day of the week. The state of our world “remembers” which day it is over a module of seven. This gives us an ED model with seven states which has a trace. The trace is that Sunday is a non-working day. Thanks to this trace we can synchronize any life which is long enough and say, for any moment of time, which day of the week it is (which state the ED model is in). Although the ED model has seven possible interpretations, the partial trace “Sunday is a non-working day” helps us capture one of these possible interpretations.

Another example in which the model has more than one interpretation: Let us have an ED model in which there is some non-deterministic transition to states  $s_1$  and  $s_2$ . Let us split the set of states of the world which correspond to  $s_1$  or  $s_2$  in two groups on the basis of various properties. We can group them in “blue” and “green” or in “big” and “small”. A simple interpretation will correspond to each splitting. These interpretations can be many (if the set is countable the subsets will be a continuum). Certainly, the more distinctions are there in the partial trace, the less interpretations will be possible, while for a full trace we will have only one possible interpretation.

## 4 Phenomena

As we said previously, phenomena are impermanent patterns. For the description of permanent patterns we used ED models. In the description of impermanent patterns we will add trigger points and terminal conditions. A trigger point will consist of a given event, probability and state of the ED model. That is, the occurrence of the given event will trigger the phenomenon with the given

probability, and that phenomenon will begin from the given state of the ED model. A terminal condition will consist only of a given event and probability: the occurrence of the given event will lead to a termination of the phenomenon with the given probability. (In the terminal conditions we have not included a state of the ED model, but we do not need to because if we wish we can always add it to the terminating event.)

**Definition 12.** A phenomenon will be the 3-tuple  $\langle Triggers, Model, Conditions \rangle$ . Here *Triggers* is the set of trigger points, *Model* is an ED model, and *Conditions* is the set of terminal conditions.

**Definition 13.** A trigger point will be the 3-tuple  $\langle e, [a, b], s \rangle$ . Here  $e$  is an event,  $[a, b]$  is a probability (probability interval) and  $s$  is state of the ED model in which the phenomenon will begin.

**Definition 14.** A terminal condition will be the 2-tuple  $\langle e, [a, b] \rangle$ . Here  $e$  is an event and  $[a, b]$  is a probability (probability interval).

**Note:** This definition implies that the trigger points and the terminating events of a phenomenon are unrelated. If we wish to relate them, we can define several different phenomena (one for each trigger point). These phenomena will have the same ED model and different terminating events.

**Definition 15.** A manifestation of a phenomenon is a segment of life which starts with a trigger point and ends in a terminal condition. In addition, the ED model pattern has to be fulfilled within that segment.

When the probability of a trigger point is 1 (e.g.  $\langle e, [1, 1], s \rangle$ ), then the occurrence of event  $e$  will always trigger a manifestation of the phenomenon. Conversely, when we have some segment of life which might be a manifestation of a phenomenon, that segment would not necessarily be a manifestation of the phenomenon since it may well be some haphazard occurrence (this may be the case only if the probability of the trigger point is lesser than 1).

## 5 Events

This paper does not provide a formal definition of the term *event*. Thus, that the formal definition of this concept remains an open issue. Instead of a formal definition of what is an *event*, we will give many examples which will provide an informal understanding of that term. Although we have not arrived at a definition of an event, we will use this term for defining more complex objects such as ED models.

We will examine two types of events: relevant and irrelevant, and three groups of relevant events: visible, invisible and semi-visible.

### 5.1 Visible events

These will be the events which depend only on the trace of life (the actions and observations). They depend on what we have already seen and on what we are going to see.

**Note:** Why do we say that an event may depend on the future? Because there will come a moment when the future becomes part of the past and because future can be predicted. Thus, a visible event will become a seen event at some later moment and can be predicted with a certain degree of confidence even before we have seen it. The future is unknown from the perspective of the current moment, but from the perspective of another moment it may already be known. For example, from the perspective of the endpoint of life, all future will already be part of the past and will be known.

Let us have the set *All* of all infinite sequences of actions and observations:

$$\dots, a_{t-3}, o_{t-2}, a_{t-1}, o_t, a_{t+1}, o_{t+2}, a_{t+3}, \dots \tag{15}$$

These sequences are infinite in both directions (towards the past and towards the future). Important here is which is the current moment (where does the zero stand). Thus, each infinite sequence can be regarded as a function:  $\mathbb{Z} \rightarrow \Sigma \cup \Omega$ .

**Definition 16.** Any subset of *All* is a visible event.

In other words, for each infinite sequence of actions and observations and for each particular moment *t* we can say whether the event occurs or not.

Now, with this definition the visible events have become too many. The members of *All* are continuum many meaning that the visible events are two to the power of continuum.

We do not have the infinite sequence of actions and observations, but only a finite segment (we have the trace of a concrete life). We want visible events to be determined only by this finite segment. Then the visible event will have three possible values: True, False or “Do not know”.

**Definition 17.** A visible event over a segment:

1. Will be True if each and every continuation of the segment to an infinite sequence results in a sequence for which the event is True;
2. Will be False if each and every continuation of the segment to an infinite sequence results in a sequence for which the event is False; and
3. Will be “Do not know” otherwise.

By reducing the visible events to segments we strongly reduced the number of these events. From two to the power of continuum the number is now continuum many. What kind of events were dropped out? Example: “Event *A* will occur infinite number of times”. If we reduce this event to a finite segment, it will

always be “Do not know” because we can continue the sequence either way. Certainly, these events are not interesting because we are interested only in those events which we can understand at a given moment.

We will not examine all visible events. Instead, we will examine only the computable ones, withal not all computable ones but only a small part of them. The most important visible events are the atomic ones:  $Ob(t)$  and  $Act(t)$ . Composite events which are composed of visible events will also be visible events.

## 5.2 Composite events

We can combine a certain number of atomic events into composite events. The conjunction of atomic events is a composite event.

If we choose to assign a name to the conjunction and convert it into an atomic formula, we will introduce an abbreviating symbol and it will be the name of the conjunction. (We will not observe a disjunction, but can observe the conjunction of negations which is the same thing.) For any composite event we can introduce an abbreviating symbol and start observing it.

We will introduce the conjunction abbreviating symbols by special heuristics which have a confidence coefficient of 100%. The form of these heuristics will be:

$$A_1(t - i_1) \& A_2(t - i_2) \& \dots \& A_n(t) \Rightarrow Abbreviation(t), \quad i_j \geq 0, i_n = 0. \quad (16)$$

We will add  $n$  more heuristics for the negations:

$$\neg A_j(t - i_j) \Rightarrow \neg Abbreviation(t) \quad (100\%) \quad (17)$$

Thus, a conjunction of atomic events will become an atomic event after we determine an abbreviating symbol which designates that conjunction.

Typically, heuristics are obtained through statistics and their confidence coefficient is less than 100%. That is why we will assume that the heuristics introduced here have been added by default.

Another composite event is “ $A$  will occur before  $B$ ”. We will express this event as follows: *exist A after t before B in our case*. This event is visible when  $A$  and  $B$  are visible. If in some infinite life events  $A$  and  $B$  do not occur after  $t$ , this composite event will be False.

## 5.3 Invisible events

Here we will discuss Invisible Single-Moment Events (ISME). Certainly, from ISMEs we can obtain other invisible events (e.g. by conjunction).

An ISME is an event which reflects the possible past and the possible future at a particular moment  $t$ . It will depend on only one state of the world. The interpretation of an ISME will be a set of states of the world.

**Note:** Why is the interpretation of an ISME a set of states of the world rather than a characteristic? Because we consider that the model of the world is fixed and for each state of the world an ISME is either True or False. If we change the model of the world then some of the states of the new world may correspond to a *belief* of states of the previous world. Once we change the world we will also change the interpretation. For each ISME we can choose a model of the world in which the interpretation of this ISME is a set.

The interpretation explains the meaning of an ISME in a concrete world, but we are going to describe ISMEs in any world. Therefore, we will look at the things from the perspective of the agent, not from the perspective of the world.

The possible future can be presented as the tree of all possible developments. We will assume that the possible past can also be presented as a tree. Then the possible past and the possible future can be presented as a tuple of trees. Let  $All2$  be the set of all such tuples of trees.

**Definition 18.** Each subset of  $All2$  will be an ISME.

According to this definition there are too many ISMEs. We will describe six ISME types which are important to us.

### 5.3.1 Always occurs

An example of this type of event is “Event  $A$  will always occur in future”. (Likewise, for the past: “Event  $A$  has always occurred”). We would like to draw some borderline. That is why we will generalize this event to the following: “In the future, event  $A$  will occur always until the moment at which event  $B$  occurs”. The concise expression of this event will be  $A$  from  $t$  until  $B$ .

It makes sense to try and find some heuristics which predict this ISME. For this purpose we will use statistics over lived experience and will select some conjunction which is a candidate for antecedent of the heuristic. Let this conjunction be True  $m$  times and let us assume that in  $k$  of these times  $A$  from  $t$  until  $B$  has been True. Then  $k/m$  appears a good candidate for a confidence coefficient. This candidate however does not reflect the specificity that at low  $m$  values the confidence is also low, so it is better to have  $(k - 1)/m$ . In this case the coefficient will be below 100% but will tend to 100% when  $k = m$  and when  $m$  tends to infinity.

Let us assume that in our conjunction there is an invisible event. We will also assume that on the basis of certain heuristics this invisible event is True.

We can obtain this information from the guesses  $G_t$ . If there are several guesses for the invisible event, then let  $R$  be the highest confidence coefficient possible. In this case we will insert in the counters (in which we accumulate  $k$  and  $m$ )  $R$  instead of 1. In this way we acknowledge that the antecedent may not always be True.

### 5.3.2 Will occur

This type is similar to the event “ $A$  always occurs”, but cannot be obtained from it by two negations because we will be left with the event “ $A$  may occur”. The latter means that  $A$  will occur in some paths of the possible future, but we want  $A$  to occur every time (in each possible life).

Accordingly, we will introduce a new invisible event: “ $A$  occurs before  $B$ ” and will express it in the following way: *exist A after t before B in all cases*.

Again, we can use statistics in order to find some heuristics which can predict this event. We should add here that the predictions should include the case “Do not know”. This is the case when life ends before event  $A$  or  $B$  has occurred, meaning that we do not know which one ( $A$  or  $B$ ) would occur if life goes on.

There is some difference between “*in our case*” and “*in all cases*”. The first means that the event will occur in our life, while the latter means that the event will occur in each possible continuation of the future. The event “*exist A after t before B in all cases*” is invisible even if  $A$  and  $B$  are visible.

### 5.3.3 Will occur with probability $p$

The event “ $A$  occurs between  $t$  and  $B$  with probability  $p$ ” is different from the previous two events. The two events defined above have the properties “retain 0” and “retain 1”. When an event “retains 0” this means that when it is False in two states, it is False in any *belief* made of these two states. By analogy, the same applies to the property “retain 1”.

The event at hand can be False in two states, but can be True in a *belief* made of these states. For example, let in these two states  $A$  occur with probabilities  $1/4$  and  $3/4$ . Then from the two states we can make a *belief* in which  $A$  occurs with probability  $1/2$ .

### 5.3.4 Feasible event

The event “ $A$  is feasible” sits between the events “ $A$  may occur” and “ $A$  will occur”.



“ $A$  is feasible” means that there is some algorithm  $P$  such that when the agent executes  $P$  then  $A$  will occur. An example of algorithm is  $Act(t + 1) = a$ . It goes without saying that the algorithm can be much more complicated.

How can we find some heuristic for “ $A$  is feasible”? Again, let us select an conjunction which is a suitable candidate for antecedent of the heuristic. Then we will consider different cases. For each case we will find some algorithm  $P$  and an heuristic which tells us that if the antecedent is in place and if we execute algorithm  $P$ , then  $A$  will occur. If we cover all cases we can conclude that if our antecedent is in place then  $A$  is feasible.

### 5.3.5 Test event

Test events are discussed in detail in [4].

An example of a test event is “If we press the door handle, the door will open”. We can call this event “The door is not locked”. The value of the test event does not depend on whether we have made the test. Thus, the door may be locked regardless of whether we checked this or not.

### 5.3.6 ED model state

Let us have some ED model and try to find the interpretation of that model. For each state  $S_i$  of the ED model we will add the event  $E_i$  which will be True when the ED model is in state  $S_i$ . We will try to find a simple interpretation, i.e. we assume that  $E_i$  is an ISME.

For each arrow of the ED model we will add one heuristic:

$$E_i(t) \ \& \ A(t) \Rightarrow E_j(t + 1) \quad (\textit{per cent}) \tag{18}$$

where  $A$  is an observed event and the arrow is over that event from  $S_i$  to  $S_j$ .

We will describe the partial trace in a similar way. For each distinction we will add the following heuristic:

$$E_i(t) \Rightarrow A(t) \quad (\textit{per cent}) \tag{19}$$

where  $S_i$  is the state,  $A$  is the distinctive event and the per cent reflects the probability.

With these heuristics we can guess the current state of the ED model and then use the trace in order to predict the additional events.

## 5.4 Semi-visible events

We need to add semi-visible events because of incorrect moves. In order to realize that a move is incorrect we have to try to play it. Accordingly, these events will sit between the visible and the invisible ones. These will be events that we will see if we watch for them.

We will introduce an atomic event (*Correct*) and two heuristics which define that event.

$$Act(t)=a \Rightarrow Correct(a, t) \quad (100\%) \quad (20)$$

This heuristic tells that if at moment  $t$  we executed successfully action  $a$ , then action  $a$  was correct at that moment. The next heuristic tells us that if at moment  $t$  we tried unsuccessfully to execute action  $a$ , then action  $a$  was incorrect at  $t$ .

$$UnsuccessfulTry(a, t) \Rightarrow \neg Correct(a, t) \quad (100\%) \quad (21)$$

The guess *UnsuccessfulTry*( $a, t$ ) will be a different guess. Instead of deriving it through heuristics we will add it in  $g_t$  by default at each unsuccessful attempt to execute action  $a$  at  $t$ . This is the way to add information about unsuccessful attempts.

Why do we add the information about unsuccessful attempts to the guesses? We assume that at any moment the world knows which move is correct and which is incorrect, but does not know which incorrect moves the agent has tried to play. This means that the information about unsuccessful attempts will be visible only by the agent.

The two default heuristics which we added have a confidence coefficient of 100%, but there may be other statistics-derived heuristics which also tell us whether a move is incorrect. It makes sense to assume that when the agent is trained she will come to know which move is correct even without trying to play it. This training will come from the additional heuristics which the agent will derive through statistics.

## 5.5 Irrelevant events

In addition to the relevant events there will be irrelevant ones. These are events which we will ignore and will not bother to understand.

Here are three examples of irrelevant events:

1. “Which life are we living now?” Given that our lived experience is a sequence of many lives, we can ask the question “Which is the serial number of the life we are living now?”

We can assume that the behavior of the world will not change across lives, but it makes sense to assume that in her first life the agent will make certain errors due to lack of experience which she will avoid in the next lives. Nevertheless, we will regard this question as irrelevant and will therefore ignore it.

2. “How old am I and how much time is left for me?” We will ignore this question, too. In other words, we will ignore the parameters  $t$  and  $k - t$ . Here is why we ignore this question: We imagine that our agent is eternally young and does not get older as life goes on. As concerns  $k - t$ , nobody knows how much time is left for them to live.

Later we will define an extended model where  $t$  and  $k - t$  matter, but we are only interested in events which if True in one life will remain True in any forward or backward extension of that life. Hence we are only interested in events which do not depend on  $t$  and  $k - t$ .

3. We will also regard as irrelevant the events which produce unimportant information. In Section 3.7 we already provided an example in which it does not matter whether the count of occurrences of an event adds up to an even or odd number.

## 5.6 Related work

What is an *event* and who introduced the term *event* in AI? The term was first introduced by Xi-Ren Cao. He introduced *events* in his paper [5].

Subsequently, in 2008 Xi-Ren Cao and Junyu Zhang provided a definition of the term *event* in their paper [6]. That definition is not good enough, because it depends on the model and because it suggests that the last state in which the model has been is remembered, but this is not something worth remembering. This is how their definition of *event* looks like:

$$E \subseteq S \times S, \text{ where } S \text{ is a specifically chosen model.} \quad (22)$$

$$E = \{ \langle s_{i-1}, s_i \rangle \mid \text{if } E \text{ happens at moment } i \} \quad (23)$$

Why does this definition depend on the model? Because Cao and Zhang assume that there is only one model of the world, which is not the case. In [3] we prove that the world has many models and furthermore that there is a Minimal and a Maximal model of the world. *Minimal* and *maximal* in this

case refer to what the states of the model know about the past and about the future.

What should the definition be, then?

$$E \subseteq S, \text{ where } S \text{ is an arbitrary model.} \quad (24)$$

$$E = \{ s_i \mid \text{if } E \text{ happens at moment } i \} \quad (25)$$

In the various models the state can “remember” a larger or a smaller piece of the past. In Xi-Ren Cao’s definition the state remembers the last state in which the model has been. As we mentioned already, this is not something worth remembering. If we decide to remember something, it would be better to remember the last action of the agent. This will make it obvious that the agent’s actions are events.

While the definition in [6] is not perfect and needs to be improved, this does nowise diminish the credit owed to Xi-Ren Cao because he was the first to realize that observing the actions only is not enough and therefore we should generalize to the wider class of events.

Indeed actions tell us everything, but they provide too much information. Where a model monitors each and every action, it is flooded and overwhelmed by an excessive flow of information. By generalizing actions to events we can reduce the inbound information and stay only with the “important” things.

**Note:** In MDP, observations are not among the monitored events, but they are detected by trace. The trace specifies the current state and thus the observations are taken into account.

**Note:** The term *event* is used in many papers, but usually with another meaning. In Lamperti, Zanella and Zhao [7] for example, *event* is meant as an *observation*. While an observation is a special case of an event, for us *event* is a more general concept.

## 6 The world

Before creating a language for description of worlds we must be clear about what is a world. We will define the world through its perfect models. The idea is the following: Imagine that you have a painting and a perfect copy of that painting. The copy is so perfect that one cannot tell the copy from the original. In this case we can assume that the original painting and its copy are the same thing.

The world will consist of a model (which will be a Simple MDP) and an initial *belief* which will show us where life is expected to start from. We will

want the model to be both perfect (so it cannot be improved any further) and minimal.

While the minimal model is not unique and even the perfect minimal model is not unique, for every pair of two perfect minimal models we can say that the states of the first model can be expressed as *beliefs* of the states of the second model.

### 6.1 The minimal model

A minimal model is one the states of which do not know any redundant facts. That is, they do not remember anything redundant about the past and anything redundant about the future. A redundant (unnecessary) fact of the past is one which does not have any impact on the future. If it does not affect the future, is it worth remembering? A redundant fact about the future is one which does not depend on the past (i.e. cannot be predicted from the past). Does it make sense to have the model know unnecessary (redundant) things about the future? For example, when you receive a letter but have yet opened it, does the world know what does the letter say? We will assume that the past can nowise help us predict the content of the letter. In this case the content of the letter is something redundant for the world, so the world does not need to know it. The world can decide what the content of the letter is only when you open it.

Of course you assume that you live in a real world and the world knows what is in the letter even before you open it. But imagine that you live in a computer program (as in the movie *The Matrix*) and the world will decide what must happen in the last minute. Thus, the world will decide what is in the letter not at the time it was written but as late as when you open it.

The content of the letter is something redundant for the world, but not for the agent. This fact is important for the agent because it is related to its future, but will be unpredictable by the agent because it is unrelated to the past. Although the fact is unpredictable, the agent will try to predict it because there is nothing to tell her that the content is unpredictable and therefore her attempts will be futile.

### 6.2 The perfect model

As we said, the states of a minimal model do not know anything redundant. Now we say that the states of a perfect model know everything useful. Therefore,

the states of a perfect minimal model know exactly what needs to be known (everything useful and nothing redundant).

There are many ways to say that a model is perfect:

1. The model has the Markov property.
2. The future depends only on our state of departure and does not depend on how we have arrived at that state.
3. The model cannot be improved any further into another model capable of making a better prediction of the future on the basis of past.
4. No state can be divided in two states such that the two new states have different past and different future. A different past means that we would be able to distinguish the states on the basis of what happened. To distinguish them means to distinguish with certainty or to say that the one is more probable than the other. It is not mandatory for each possible past to distinguish the two states. Just one possible past that can distinguish the two states is sufficient. By analogy we can say what it means when two states have a different future.

These definitions of a perfect model are all equivalent. The state of the perfect model need not be split in two because it knows everything useful. If we split the state in two on the basis of the past, the two new states will know something more about the past, but that thing will not be useful and therefore the new states will be bound to the same future.

**Question:** If we pick randomly some model, will it be a perfect model of some world? The answer is yes. Every model is a perfect model of some world. Every model describes some world and if we assume that it is the best possible description and that description cannot be improved any further, then the model is perfect. Of course there are infinitely many worlds for which that model provides some partial description which can be improved. The main question is: “Have we found the perfect model exactly of the world we are trying to describe?” The answer to this question is that we do not know. All we know about the world is our lived experience. This experience can occur in infinitely many worlds. Certainly, we are looking for the simplest model that matches our experience (this principle is known as the Occam’s razor).

In fact, we cannot even be sure that the model built on the basis of our lived experience is correct. There are some probabilities in the model which we determine by statistics. When the collected statistics are limited these probabilities may not be accurate, but if we assume that the statistics are sufficient then the only question left would be whether the model is perfect.

We will not try to find a perfect model of the world. All we need is a sufficiently good model which works for us. In [8] we explained why trying to find a perfect model of the world is an overly ambitious goal. Nevertheless, we

will assume that a perfect model of the world does exist and that model provides the definition of the world.

Furthermore, we are not looking for a minimal model of the world because it would make us admit that certain events are unpredictable. As a rule, we will try to predict any event even though we assume that the result from certain events (such as casting a dice) is unpredictable.

### 6.3 The basic model

**Definition 19.** A basic model of the world is every perfect and minimal model of the world.

We will assume that each world has a basic model which provides the definition of that world. A model can be the basic (perfect) one of a certain world, but for another world it may be just an ordinary (imperfect) model.

The states of the basic model know everything useful both about the past (everything that could have happened) and about the future (everything that can happen going forward).

**Note:** Various authors may use different terminology. Thus, the term used in [9] is “game” instead of “world”. Likewise, [9] uses “Game Description Language” instead of “Language for Description of Worlds” as well as “Imperfect Information” instead of “Partial Observability”. We can imagine the world as a game and, vice versa, a game can be thought of as a world in its own right. The Mystery group uses this idea in their song “The World is a Game”.

### 6.4 Simple MDP

Worlds are usually described using Markov decision process (MDP). We will simplify the MDP and create a Simple MDP.

Why does MDP need simplification? Why do we claim that there is a degree of needless complication in the MDP? There are two reasons for this.

First, MDP conceals the fact that besides being a model of the world it also models the agent. Indeed, the world and the agent form a single system, so one cannot describe the first without describing the latter. In other words, MDP describes not only the behavior of the world but also the behavior of the agent. If you look closely at the MDP you will see that according to MDP the agent does whatever she wants. Doing whatever you want is also a kind of behavior, though it is the most liberal behavior possible. What is the probability for the agent to choose a particular action? We do not know – the agent is free to do whatever she likes, so the probability is in the interval  $[0, 1]$ .

Second, the MDP restrains the world and forces it to follow some fixed policy. I.e., in MDP the agent is unrestrained and free to do anything, while the world is restrained to a fixed policy.

## 6.5 Fixed policy

What is a fixed policy? For example, this is the case when in a given situation you always turn left. Even more so, this is an extreme policy because *always left* and *always right* are two extreme possibilities. But, in a scenario when you toss a coin and if it falls heads then you turn left you will be executing a fixed policy with probability  $1/2$  (exactly  $1/2$ ). If the probability is different then you are executing a different fixed policy. If you turn left or right as you like, then you will not execute a fixed policy and instead will turn left with probability in the interval  $[0, 1]$ .

Do you have free will (freedom of choice) when executing a fixed policy? The answer is *no*. When you toss a coin, it is the coin which makes the decision, not you. Thus, in MDP the agent enjoys absolute freedom when choosing her move. In MDP the agent is not restrained by anything, while the world is absolutely restrained as it is forced to execute a fixed policy.

## 6.6 Extreme policy

An extreme policy is a fixed policy in which every probability is set to its minimum or to its maximum value.

How can we select an extreme policy? If the probabilities are determined by the intervals  $[a_i, b_i]$ , we will select the extreme value of one of these intervals (i.e. we select either  $a$  or  $b$ ). If we select  $b$ , this may shorten the remaining intervals. Then we will select one of these remaining intervals (probably shortened) and so forth. In this way we select a fixed policy where every probability is extreme (i.e. it cannot be increased any further or it cannot be reduced).

When we make a reference to policy in Simple MDP, we will mean a policy shared by the agent and the world. To put it another way, the two of them have colluded to implement two policies, the combination of which is the policy of the Simple MDP. In the MDP, references to a policy are in fact references only to the policy of the agent because in the MDP the world is bound to a fixed policy and has nowhere else to go. Then, the agent's policy in the MDP is in fact an extreme policy (each one with probability either 0 or 1). What makes extreme policies sufficient in the MDP? When we pursue a certain goal and have to decide *left or right*, typically there are three scenarios: 1) our goal



will be better served when we go left; 2) our goal will be better served when we go right; and 3) it does not matter whether we go left or right. If we assume that in the third scenario we choose to go left, then we end up with an extreme policy. This makes extreme policies sufficient in the MDP. Certainly, this does not apply to each goal. If the goal is diversity (tour around as many states as possible), then we should alternate left and right.

The good thing about extreme policies is that when the states are finitely many, then the extreme policies are also finitely many, while the fixed policies usually are continuum many.

## 6.7 From MDP to Simple MDP

The MDP can help us describe only part of the worlds. These are some very peculiar worlds where nature does not have free will as it is forced to execute a fixed policy. If an agent with free will lives in such a world, the world cannot be described by MDP. Let us take the world of the chess game in which there is a second agent who plays against the protagonist. Let us assume that the second agent is not deterministic (she is not bound to play some extreme policy). Suppose that she even does not play a fixed policy. So, we let the second agent loose and she can play whatever she wants. This type of world cannot be presented by the MDP, but we can describe it using the Simple MDP. The difference between the two models is that instead of exact probabilities, in Simple MDPs we operate with probability intervals.

The MDP is needlessly complicated because it obscures the fact that the probability of the agent's actions is actually the probability interval  $[0, 1]$ . While the arrows over actions come with probabilities, rather than determining the probability of the action these probabilities determine indirectly the probability of the observation. (The observation has some influence on what the next state of the MDP model will be, but that influence is very subtle and indirect. The arrows define, in a non-deterministic manner, several possible states and some of these possibilities are dropped out on the basis of the observation. With Simple MDP things are much more straightforward because the Simple MDP includes arrows over observations which directly indicate the influence of the observation on what the next state will be.)

In MDP the state has past, present and future. That is, something has happened before the state, something is happening within the state and something will happen after the state. The Simple MDP state has only past and future because nothing happens within the state. In the MDP state there is some observation. There are two equivalent definitions of MDP. We will name them

*monochromic* and *polychromic*. With the monochromic definition in the state there is only one possible observation, while the polychromic definition says that in each state there are several possible observations and each observation comes with a fixed probability. This means that the model in the monochromic definition is not minimal because the state “knows” exactly which observation will come up and that knowledge may not follow from the past. It may be the case that the past provides for a non-deterministic transition between several monochromic states. Minimality will occur if these several monochromic states are replaced with a single polychromic state wherein each observation comes with its probability.

## 6.8 Definition of Simple MDP

We will present the Simple MDP as a finite automaton (except for the requirement that the number of state is finite). More precisely, we will present it as a probabilistic automaton because over the arrows we will have probabilities. Even more precisely, we will present it as interval-valued probabilistic automaton because over the arrows instead of probabilities there will be probability intervals.

With the MDP there is only one type of state due to the fact that in MDP it is always the agent’s turn to make a move. The world in Simple MDP will be a game between the agent and the world. Thus, there will be two types of states of the world: 1) states in which it is the agent’s turn to make a move and 2) states in which it is the world’s turn to make a move.

Each arrow in MDP stands for one move. This move comprises some action by the agent and some response by the world. (The response of the world is the observation which the agent sees.) In Simple MDP each arrow will represent a ply (half move). This half move is either the agent’s action or the world’s response.

**Definition 20.** Simple MDP is a graph  $\langle U \cup W, A \cup O \rangle$  with two types of nodes and two types of arrows:

- $U$  are the states of the world when it is the agent’s turn to make a move (the agent will execute some action);
- $W$  are the states of the world when it is the world’s turn to make a move (the world will display some observation to the agent);
- $A$  are arrows which represent actions ( $A$ -type arrows);
- $O$  are arrows which represent observations ( $O$ -type arrows);

- If  $u \in U$ , then the arrows departing from  $u$  are  $A$ -type arrows and the arrows ending in  $u$  are  $O$ -type arrows;
- If  $w \in W$ , then the arrows departing from  $w$  are  $O$ -type arrows and the arrows ending in  $w$  are  $A$ -type arrows;
- $A \rightarrow \Sigma \times [0, 1] \times [0, 1]$ , meaning that each  $A$ -type arrow is associated with some action and a probability interval;
- $O \rightarrow \Omega \times [0, 1] \times [0, 1]$ , meaning that each  $O$ -type arrow is associated with some observation and a probability interval.

We must add something about the probability intervals. When the probability is fixed (the length of the intervals is zero), then the sum of the probabilities of the arrows which depart from one node must be 1. When the probability is not fixed, then we will consider the probability intervals (of the arrows departing from one node) as a description of a set of vectors of fixed probabilities, wherein the sum of each of these vectors is 1. This means that the intervals  $[a_i, b_i]$  describe vector  $p_i$ , where  $a_i \leq p_i \leq b_i$  and  $\sum(p_i) = 1$ . We will want this description to describe at least one probability vector (i.e. the set of the described vectors must not be the empty set). We will also want the description to be optimal (meaning that each shortening of the intervals will remove some vector from the set). In [10] there are several inequalities which must be true for the description to be both non-empty and optimal.

## 6.9 MDP as Simple MDP

How can we present MDP as a Simple MDP? In MDP all arrows are over actions, which makes them  $A$ -type arrows. Each arrow will remain in the Simple MDP, the only difference being that its probability  $p$  will be replaced with the interval  $[0, p]$ . Hence, if the agent chooses that action, then she will choose that arrow with probability  $p$  and in the opposite case will choose that arrow with probability 0. Where the action is only one (no arrows over other actions depart from the state), then  $p$  is replaced with the interval  $[p, p]$ .

How do we change states? Each state  $s$  is replaced with two states ( $w$  and  $u$ ) which belong to type  $W$  and accordingly to type  $U$ . All arrows which previously arrived at state  $s$  will now arrive at state  $w$  and the arrows which previously departed from  $s$  will now depart from  $u$ . We will add additional arrows from  $w$  to  $u$ . The number of these additional arrows will be equal to the number of possible observations  $ob$  in  $s$ . Each such arrow will be of type  $O$ , will be over

the corresponding observation  $ob$  and will be associated with the probability interval  $[p, p]$ , where  $p$  is the probability of the observation  $ob$  in  $s$ . What shall we do in the case of the monochromic definition (where each state has one and only one observation)? Then there will be only one arrow from  $w$  to  $u$  and its probability interval will be  $[1, 1]$ .

The resulting Simple MDP will describe the same world as will the MDP. If the MDP model has been perfect, then the resulting Simple MDP will also be perfect. The same applies to minimality.

So we have proven that all worlds which can be described through MDP can also be described through Simple MDP but not vice versa. In other words, Simple MDP expands the notion of *world*.

### 6.10 Initial belief

Before we can describe the world we need to add an initial state. We prefer to have a set of possible initial states instead of a single initial state. Each of these possible states comes with some probability. If this probability is fixed, we will have a structure which we will designate as fixed *belief*.

**Definition 21.** Fixed *belief*:

- $M \subseteq U \cup W$ .
- $M \rightarrow [0, 1]$ .

This means that we have a set of states and have matched each state in that set to a fixed probability. The sum of these fixed probabilities must be 1. In this case saying that  $s \notin M$  is not much different from saying that the probability of  $s$  is zero. The term *belief* has the meaning used in the MDP rather in the meaning employed in “deep belief network”. In most papers about MDP our notion of fixed *belief* is called *belief*, but here *belief* will be something more complex.

**Definition 22.** A generalized *belief* is the set of fixed *beliefs*.

For the sake of simplicity instead of generalized *belief* we will use just *belief*.

We will assume that the world is described by one Simple MDP and one initial *belief*. Which initial state do we expect to depart from? First, from the members of the initial *belief* we pick one fixed *belief*. How do we pick it? Any way we wish, here we have randomness with unknown probability. Then from the selected fixed *belief* we pick a concrete initial state with the probability given by that fixed *belief*.

The meaning of the notion *initial belief* can be inferred from lived experience and from the statistics of the world. This is the expectation about what the initial state of the world will be.

## 7 Interpretation

We already explained what is an interpretation of an ED model, but have not yet explained the interpretation of an event.

We will introduce the extended model. We will obtain it by adding the past and the future to the basic model. Thus, while the states of the basic model know what could have happened and what can happen going forward, the states of the extended model will know what exactly has happened so far and what exactly will happen from this moment onwards.

The interpretation of an event will be a set of states of the extended model. Each event will have its own probability because each state of the extended model will have its own probability.

In order to simplify the explanation, we will assume that the world is cyclic. The property of being cyclic is characterized by two things:

1. Each event that has already happened may happen again.
2. The initial *belief* is the invariant *belief*. Thus, the question “In which state do we expect to be at the initial moment?” coincides with the question “In which state do we expect to be at a random moment?”

The good thing about the cyclic world is that in it the backward probability does not depend on  $t$ . Another advantage of the cyclic world is that in it we can define a probability for each state of the basic model and that probability will not depend on  $t$  either.

### 7.1 Possible future

The possible future is a sequence of observations, actions and sets of incorrect moves. These sets are finite and that is why the possible future is a word over an alphabet composed of  $n+m+2^m$  letters.

For each state  $s$  we can describe the possible future as an infinite tree of all words which the future can produce if we start from  $s$ . Each node of the tree will be associated with some word (possible future) and some probability (probability interval).

The probability is the product of the probabilities which we have captured in the Simple MDP model when we start from  $s$  and read the word in question. (Arrows are associated with actions and observations, while  $U$ -type states are

associated with sets of incorrect moves.) If the word in question can be read in two or more ways, then the probability is equal to the sum of the various probabilities which may be obtained by the different readings.

Where the probability is an interval of probabilities we should explain how probability intervals are added and multiplied.

$$[a_1, b_1] \cdot [a_2, b_2] = [a_1 \cdot a_2, b_1 \cdot b_2] \quad (\text{Multiplication}) \quad (26)$$

$$[a_1, b_1] + [a_2, b_2] = [a_1 + a_2, \min(b_1 + b_2, 1)] \quad (\text{Addition}) \quad (27)$$

How the tree of the possible future looks like? The number of arrows over one action which depart from one node cannot be more than one, but the number of arrows over one observation can be more than one. When all moves are correct then the tree is simple and the arrows over observations cannot branch out, either. The presence of incorrect moves makes the tree more complicated because from one node there can depart several arrows over one and the same observation, but they will end in nodes which correspond to different sets of incorrect moves.

We will assume that none of the arrows is associated with a probability of zero. If there is such probability, we will remove that subtree. If we let some arrows have zero probabilities, then the tree will not be unique.

## 7.2 Possible past

We prefer to reuse the definition of the possible future and apply it to the possible past. However, we have a problem here. For each arrow we will need its backward probability, that is the probability of arriving from that arrow. (We already have the forward probability, but it means something else: how probable it is that we departed over that arrow.)

We can define the backward probability by the statistics of the world. Its value will be  $k/m$ , but  $m$  here will express the number of times we have been through the head of the arrow (not through its tail). The problem here is that in defining the forward probability we assumed that it is permanent and does not depend on  $t$  (i.e. does not depend on the step at which we are). However, when defining the backward probability it may turn out that it depends on  $t$ .

In order to calculate the backward probability, we will need the forward probability and something else, namely the probability of each state. In the initial moment we can present that probability through an initial *belief*. For each next moment we can calculate the next *belief* by using the initial *belief* and the forward probability.

The formulas for this purpose are:

$$m_i \cdot p_i = m \cdot q_i \quad (28)$$

$$q_i = \frac{m_i \cdot p_i}{m} \quad (29)$$

$$m = \sum (m_i \cdot p_i) \quad (30)$$

Here  $q_i$  are the backward probabilities of some state  $s$ . The index  $i$  runs the previous states (the states of departure of an arrow which arrives at  $s$ ). The probabilities  $p_i$  are the forward probabilities of the arrows which depart from the previous states and arrive at state  $s$ . The probabilities  $m_i$  are obtained from the previous *belief*. These say how likely it is that we have been in some previous states at the previous moment, while the probability  $m$  says how likely it is that at the next moment we will be in state  $s$ .

The basis of the first equality is that the probability of departing from a certain state over a certain arrow is equal to the probability of arriving at the next state over the same arrow. The formulas above determine the backward probabilities  $q_i$  in all cases except one. This is the case when  $m = 0$ . In this case we can assign to  $q_i$  any values we wish (as long as the sum of  $q_i$  is 1).

Thus, for each step  $t$  we obtain some *belief* and backward probabilities which depend on step  $t$ . We defined the possible future as an infinite tree, but the possible past will be much more complex. It will be a countable set of trees. At moment 0 we will not have any past (i.e. a tree with a depth of 0). The possible past at moment  $t$  will be presented as a tree with a depth of  $t$ .

Preferably, the possible past should be as simple as the possible future. For this purpose the expectation about the state in which we are must be constant (independent from step  $t$ ).

### 7.3 Past vs Future

In this paper the past and the future are absolutely symmetric. This symmetry comes from the fact that we present life as a path in a certain graph. The direction we follow does not make a difference, i.e. moving forward in the direction of the arrows is not any different from moving backward against the arrows.

At a first glance it appears that there is some difference between forward and backward probabilities, but in fact there is not any difference between these probabilities, either. The forward probabilities are constant, while the backward probabilities depend on  $t$ , but this is because we first establish the

forward probabilities and thereafter establish the backward probabilities. With the inverse approach (establish the backward probabilities first and on their basis establish the forward probabilities) the backward probabilities would be constant and the forward probabilities would depend on  $(k-t)$ . Certainly, when we walk backwards we need to have a final *belief* instead of an initial *belief*. Hence, we must depart from some final *belief*.

Later on we will introduce the notion of *cyclic world*. This will be a special case where the forward and backward probabilities do not depend on  $t$ . In this way we will demonstrate that in the interesting case the past and the future are absolutely symmetrical.

#### 7.4 Free will

We assumed that the world and the agent may have free will (freedom of choice). I.e. we suggest that the world and the agent can do whatever they like (but within certain limits). Since we describe the world through statistics, the players must make use of their freedom, otherwise the model will not be able to capture that freedom.

If the agent and the world do not exercise their freedom to the fullest extent, these statistics will produce some abridged model of the world which reflects only the actual behavior of the players.

For example, in a chess game the agent and the world have some correct moves and can play any of them. But in an actual game they will not play each and every correct move, e.g. we may reasonably suggest that they will not make stupid errors. If we use statistics to describe the chess game, the probability intervals perhaps will be shorter, i.e. the resulting model will not capture all of the possible moves. This will reflect the actual behavior of the agent and of her opponent (there are possible moves which the two of them will never play).

In chess there is one rule which says: "You cannot make a move if it enables the opponent to capture your king in the next move". We can say that such a move would be a stupid error and the rules of the game prohibit stupid errors. So, this rule reduces the number of correct moves. If we prohibit other stupid errors we will end up with a more abridged model in which the behavior of the world and agent will be even more restricted.

#### 7.5 Invariant *belief*

We wish to have an initial *belief* such that the next *belief* is the same and every next *belief* is the same. I.e. we wish the initial *belief* to be an invariant



one (an invariant point).

We will assume that the probability interval of each arrow is different from 0 (i.e. the interval can be  $[0, p]$  but cannot be  $[0, 0]$ ).

**Definition 23.** A trap is a group of states which one can enter but can never exit.

If there are not any traps in the world, then each connected component will be a strongly connected graph. Let us assume that in each connected component we have selected one strongly connected component to be the main one (let's call it the kernel of the connected component). There are not any traps in the world if and only if there are not any inflows and outflows.

**Definition 24.** Inflow is a group of states from which there is a way into the kernel, but no way back.

**Definition 25.** Outflow is a group of states to which there is a way from the kernel, but from which there is no way back.

For an invariant *belief* to exist, the probability of coming from an inflow has to be zero. Otherwise, the probability of being in the inflow will decrease, but it has to be constant. Therefore, for invariant *belief* to exist, there must be no inflows. Analogically for outflows.

## 7.6 Cyclic world

Let us first define the concept of compact world:

**Definition 26.** A compact world is a world without any inflows and outflows.

**Definition 27.** A cyclic world is a world which satisfies the following three conditions:

1. There is a path between any pair of states (strongly connected graph);
2. The probability interval of each arrow is different from 0; and
3. The initial *belief* of the world is the set of all invariant fixed *beliefs*.

The first condition means that the cyclic world is compact and has one connected component. The second condition ensures that we can go into every state with some probability if there is some support from the world and the agent (through their free will). The third condition means that the following two questions have the same answer:

1. Which state do we expect to depart from?
2. In which state do we expect to be after some very long life, regardless of our state of departure?

We will assume that the world is compact and even cyclic. The good thing about the compact world is that the backward probabilities are unambiguously

defined so that the possible past is simple (it is one tree). The good thing about the cyclic world is that we can define unambiguously the invariant *belief* and thereby provide a more simple definition of the extended model. In the opposite case the expectation as to which particular state we are in will depend on the connected component and on the way in which the initial *belief* distributed the probability over different connected components.

In [11] we assumed that our objective is to find some AI which will work well in worlds that are free from fatal errors. The assumption in [11] was that if some AI works well in this kind of worlds, it would work well in any world. The same assumption can be made in respect of the cyclic worlds.

Can we say that a cyclic world is free from fatal errors? In the next sections we will see that if a world is cyclic and is not hostile, then any error would not be fatal. What do we mean by a hostile world? We said that the world enjoys some freedom of choice so it can choose its strictly defined policy from a pool of strictly defined policies. A hostile world can intentionally choose a policy which works against the agent. Then the agent may end up in some concavity without being able to crawl out of it because the world would not let her get out. If the world is not hostile (i.e. if it is impartial or even good-minded and helpful), then the agent would be able to get of any concavity it falls in so there will not be any fatal errors.

### 7.7 A unique *belief*

When is the invariant fixed *belief* unique? Let's look at the more simple case when we have a fixed policy (i.e. the probabilities  $p_i$  are fixed). Let also the probabilities  $p_i$  be different from zero. And finally, let the world be cyclic.

In this case we would have a unique invariant *belief* (that belief will be fixed). We can calculate it by solving a system of equations. Another way to obtain this fixed *belief* is through the statistics of the world. We can pick some very long life. It does not matter which is the initial state. Let then  $m$  be the probability of us being in state  $s$ . That probability will tend to the value given to us by the unique *belief*. Here we do not mean the probability at moment  $t$ , but the average probability. The reason is that the probability may have a waveform pattern and at the moment  $t$  it may be divergent, while the average probability is convergent.

So we obtained an invariant *belief* which answers the natural question: "Which state do you believe the world is in?"

Let us now the world be compact, but not cyclic and let there be two connected components. Then the invariant *belief* will not be unique because we can

distribute the probability of the two connected components in any ratio (e.g. 1:1 or 1:2).

Now let the world be cyclic, but with a fixed policy where some probabilities are equal to zero. Then this policy will divide the world in concavities with flat bottoms and surroundings around the concavities. Once we hit the bottom of a concavity we will never be able to get out of it. If we depart from some surrounding around a concavity we will end up in the bottom of that or of some other concavity. In this case the invariant *belief* will not be unique because the probability in the surroundings around the concavities will be zero (it will be distributed across the bottoms of the concavities). The problem is that the probability can be distributed in any ratios across the bottoms.

How can we in this case construct a unique *belief*? Let us look at the limit where the probabilities  $p_i$  are different from zero, but tend to the ones we have. In this way we will find the natural distribution of the probabilities over the various bottoms. This is how we can find a unique invariant *belief* when the policy is fixed.

If we take the set of all fixed policies (which fall within the probability intervals) we will obtain a set of fixed *beliefs*. That will be the generalized *belief* which answers the question: “Which state do I believe I am?”

For each state of basic model the generalized *belief* will return one probability interval. That interval will tell us how likely it is that we are in a certain state of the basic mode (the probability of us being in that state). Let’s designate this interval with the function *expected*.

### 7.8 Extended model

We said that the states of the basic model know everything useful and nothing redundant. In the extended model we will add one piece of redundant information. We will add the entire life (the trace and the backbone). This information is “redundant” because the state of the basic model “knows” what is the possible past and the possible future, but does not know what exactly happened in the past and what exactly will happen in the future.

The extended model will have a set of states  $S'$ .

$$S' = \{ \langle t, L \rangle \mid 0 \leq t \leq k, L \text{ is possible life} \} \tag{31}$$

Here  $k = |L|$ . What do we mean by “possible life”? This is any finite path in the graph which describes the basic model. (Any state can be the initial state because of our assumption that the world is cyclic.)

The set  $S'$  will be the *set of moments* because each moment in each possible life is a member of that set.

If we present the extended model as a graph it will have a very simple structure composed of non-intersecting threads. Each thread will be associated with one possible life.

The state of the extended model will know what exactly has happened and what exactly is going to happen, however, this does not mean that if we know the past and the model we will be able to foretell the future. While it is true that the state “knows” everything, we cannot know which particular state we are in.

When  $t = 0$ , then the initial *belief* defines the particular state of the extended model we are in. We will obtain the new initial *belief* from the set  $S'_0$  by adding to each state the probability  $expected(\langle 0, L \rangle)$  (these definitions are provided below).

$$S'_0 = \{ \langle 0, L \rangle \mid L \text{ is possible life} \} \quad (32)$$

Thus, the new initial *belief* will contain the initial moment of each possible life. These members are many because even if fix the initial state  $u_0$ , then the members of the new initial *belief* will again be countably many, because the possible lives starting from  $u_0$  are countably many (due to the variety of lengths and branches). This means that state  $\langle 0, L \rangle$  “knows” everything about the future, but we do not know it because we have no way to know that we are exactly in state  $\langle 0, L \rangle$ .

We will now extend the function *expected*, which is defined for the states of the basic model, by defining it for the moments (i.e. the states of the extended model). For this purpose will take the probability of life  $L$  and will divide it by the number of moments in  $L$ .

We will want the sum of the probabilities of all lives to be 1 and that is why we will assume that the life can be of length  $k$  with a probability of  $(1 - \lambda) \cdot \lambda^k$ . (So we distribute 1 across all possible lengths of the life.) We will select some coefficient  $\lambda$ , however, the way we choose that coefficient is not important because we assumed that events are truth-preserving under life extension. (The formula for *expected* will depend on  $L$  but not on  $t$  because the probability is distributed evenly across all moments of life  $L$ .)

$$expected(\langle t, L \rangle) = \frac{expected(s_0) \cdot p_1 \cdot p_2 \cdot \dots \cdot p_k}{k + 1} \cdot (1 - \lambda) \cdot \lambda^k \quad (33)$$

Here  $k = |L|$ ,  $s_0$  is the first state of life  $L$  and  $p_i$  is the probability of the

arrow from  $s_{i-1}$  to  $s_i$ . Let  $q_i$  be the backward probability of the arrow from  $s_{i-1}$  to  $s_i$ . Then we have three ways to express the probability that some life of length  $k$  is exactly life  $L$ :

$$expected(s_0) \cdot p_1 \cdot p_2 \cdot \dots \cdot p_k \tag{34}$$

$$q_1 \cdot q_2 \cdot \dots \cdot q_k \cdot expected(s_k) \tag{35}$$

$$q_1 \cdot \dots \cdot q_t \cdot expected(s_t) \cdot p_{t+1} \cdot \dots \cdot p_k \tag{36}$$

That is, we can start from the probability of any state along the backbone of life and multiply it by the probability of transitioning to the next (or the previous) state.

**Definition 28.** The interpretation of event  $A$  is the tuple of states  $P$  and  $Q$ , where  $P$  are the moments at which  $A$  occurs,  $Q$  are the moments at which  $A$  does not occur, and the remaining moments are those for which we do not know whether  $A$  occurs or not.

**Definition 29.** The probability of event  $A$  is:

$$expected(A) = \frac{expected(P)}{expected(P) + expected(Q)} \tag{37}$$

This means that the probability is determined by the moments at which we know what the value of the event is.

We assumed that events are truth-preserving under life extension, which translates in the assumption that the sets  $P$  and  $Q$  are closed under the “life extension” operation.

Our definition of event depends on the selected basic model. A different basic model would come with different  $\langle P, Q \rangle$  tuple. Nevertheless, the event will be the same because it will have the same probability and will behave in the same way versus the trace of life.

## 7.9 MDP as an ED model

As an application of the above we will prove that MDP is a special case of an ED model. With MDP models the observed events are the agent’s actions:

$$act(t) = a_i, \quad 1 \leq i \leq m \tag{38}$$

These events do not intersect so we do not need any rules for the reconciliation of collisions. The additional events (those which participate in the partial trace) are the current observations:

$$ob(t) = o_i \vee ob(t+1) = o_i, \quad 1 \leq i \leq n \quad (39)$$

Here it is not clear whether  $t$  is an even or odd number and whether the current observation has just occurred or will occur at the next step. For this reason the event is defined by disjunction.

### 7.9.1 Fully observable Markov decision process

Let's first look at the Fully observable MDP (FOMDP). In this case the topology of the ED model is a simple one. The model has  $n$  states (the number of possible observations) and there is an arrow over each action between each pair of states. The trace of the ED model is clear and nice. In each state there is one current observation which occurs with a probability of 1 (each state is identified unambiguously by the current observation). The only thing we need to do in order to define the FOMDP model are the probabilities over the arrows:

$$p_{iaj} = \text{expected}(ob(t+2) = o_j \mid ob(t) = o_i \ \& \ act(t+1) = a) \quad (40)$$

This formula says that the next observation is  $o_j$ , given that the previous observation has been  $o_i$  and the last action has been  $a$ . Here we should explain what does it mean that an event is under a condition. This means that the event will be True or False when the condition is fulfilled and "Do not know" when the condition is not fulfilled.

The function *expected* can return either some strictly specified probability or some probability interval. When *expected* returns strictly defined probabilities in (40), we will obtain a classic FOMDP definition. Certainly, nothing prevents us from generalizing the FOMDP definition and thus allow for transitions the probability of which will be probability intervals.

**Note:** The FOMDP has a single unambiguous interpretation and it is defined by the event "What is the current observation?" We know what is the current observation for any moment of life except for the first moment (moment zero). Not knowing the current observation at the first moment is not a problem because the interpretation need not be defined for a finite life. It must be defined over any infinite life, and infinite lives do not have a first moment.

Thus we have obtained a model with  $n$  states and a transition probability matrix  $p_{iaj}$ . In literature this model is referred to as Full Observable MDP. The question which comes with it is whether the model satisfies the Markov property or in other words is the model perfect (the best possible model which cannot be improved any further).

On one hand, we should be happy if the model satisfies the Markov property because this tells us that we have found the best solution. On the other hand, we should not be so happy because a perfect model means that we have hit a stonewall beyond which it is not possible to improve the model any further.

Accordingly, for each model we will assume that while there is chance for the model to satisfy the Markov property, it is more likely that the model does not. In other words, we will imagine that it might be the best possible model for this world, but in all probability there are better models out there.

### 7.9.2 Partially observable Markov decision process

This is the next model we will deal with. Again, its topology is simple. We have a random number of states and arrows over each action between each pair of states. We have probabilities over the arrows, but also probabilities over the trace. For each state of the ED model we have  $n$  distinctions. In each state, each observation will be observed with some probability and, in the general case, that probability will be different from the average one.

The question is whether there is any interpretation of the so-described Partially observable MDP (POMDP). There may or may not be an interpretation. (If the model is inadequate, then there will not be any interpretation.) Therefore we will examine the set of possible interpretations and for each interpretation we will determine the probabilities over the arrows and over the trace.

We will examine the unambiguous interpretations. Thus, we will partition the set  $S'$  into equivalence classes ( $S'$  is the set of moments). After partitioning  $S'$  each state of the ED model will correspond to one equivalence class.

**Note:** The equivalence relation  $R$  is not fully random because  $R$  must retain the transition over observations. This means that for each odd-numbered moment  $\langle t, L \rangle$  that moment and the next moment  $\langle t + 1, L \rangle$  must be in the same equivalence class. A sufficient condition for retaining the transition over observations is to define  $R$  on the basis of some equivalence relation over the set  $O$  (the set of arrows over observations).

Let events  $C_i$  are those which determine the equivalence classes of  $R$ .

$$p_{iaj} = \text{expected}( C_j(t+2) \mid C_i(t) \ \& \ act(t+1)=a ) \quad (41)$$

Thus, the next state is  $j$  if the previous state was  $i$  and the last action has been  $a$ .

What will be the POMDP trace? For each state  $i$  we have  $n$  possible observations and each one of these observations occurs with some probability  $q_{ij}$ .

$$q_{ij} = \text{expected}(ob(t)=o_j \mid C_i(t)) \quad (42)$$

As we said already, the function *expected* can return either some strictly specified probability or some probability interval. If we aim to obtain a classic POMDP definition we have to assume that  $p_{iaj}$  and  $q_{ij}$  are strictly specified probabilities.

**Note:** If we stay with the standard definition of Partially observable MDP, then the only possible trace of the model will be the probabilities  $q_{ij}$ . However, we will assume that there may be other traces, i.e. other distinctions, which are characteristic of a given state  $C_i$ . For example, it may be the case that a certain action can never occur in that state. Another possible distinction is that a certain conjunction of events is impossible, even though the individual events in the conjunction are perfectly possible.

**Note:** The following question arises: “Is there a POMDP which is a perfect model of the world (a POMDP which satisfies the Markov property)?” We have a plethora of POMDPs (for example there is one POMDP for each equivalence relation over  $O$ ). But, among all these models, is there a perfect model? The answer is “Yes, there is.” This is how we will obtain that perfect model from the basic model: We will take all arrows in  $O$  and for each arrow we will create the event “Walk over this arrow”. This event will be True at each odd-number moment when we depart over the arrow as well as at the next moment. The so-obtained POMDP model will be perfect as a sequel the perfection of the basic model. The model will not be minimal, but if we put together all arrows which have the same beginning and the same end, then we will obtain a perfect and minimal POMDP (as a sequel of the perfection and minimality of the basic model).

## 7.10 Related work

The idea which drives the creation of ED models is that the states of the world are far too many so we should reduce them in order to make the model more understandable. The same idea is embodied by Wang, Joshi and Khardon in [12]. Their paper introduces the concept Relational MDPs, which essentially are a special case of the ED models.

We said that any equivalence relation (which partitions the set of arrows) can be used for creating a POMDP. Wang, Joshi and Khardon propose that certain events be selected for a Relational MDP. (Instead of events they refer to predicates, actions and rewards.) These events are then used to define the



equivalence relation (two states are equivalent if the same events from the pool of selected events occur in these two states). A POMPD is defined over the quotient set of that relation and this is the wanted Relational MDP.

**Note:** The definition of [12] does not use a quotient set, but refers to a family of MDPs. If each of these MDPs has only one state from each equivalence class, then it will have the structure of the quotient set, but typically this is not the case. Therefore, it is better to define Relational MDPs through the quotient set.

How does a Relational MDP look like? We will explore three scenarios:

1. Let the input consist of 10 bits. Let these 10 bits be the ten events which have been selected and used for the creation of the Relation MDP. Then the resulting Relational MDP coincided with the FOMDP.

2. Now let only 5 of these 10 bits be selected events. This will give us a Relational MDP with less states than the FOMDP states.

3. Let us now have the selected 10 input bits (events) plus 5 other events which are invisible (do not follow from the input). Let's take these 15 events and create a Relational MDP. Now the Relational MDP will have more states than the FOMDP.

## 8 Conclusion

In the first part of the paper we described the key concepts which we will need to create the language for description of worlds we are looking for. In the second part we will show how quite a complex world can be described and will see that the task can be achieved by a fairly simple description. By using the description of the complex world, we will create our language for description of worlds. In other words, our approach will be to solve the general task by exploring a concrete example: We will create the language for description of worlds on the basis of a particular world and its description.

## References

- [1] A. M. Turing, On computable numbers, with an application to the Entscheidungsproblem, *Proceedings of the London Mathematical Society*, 42:230–265, 1937.
- [2] D. Dobrev, Event-Driven Models, *International Journal “Information Models and Analyses”*, 8(1):23–58, 2019.

- [3] D. Dobrev, Minimal and Maximal Models in Reinforcement Learning, *International Journal “Information Theories and Applications”*, 26(3):268–284, 2019.
- [4] D. Dobrev, Incorrect Moves and Testable States, *International Journal “Information Theories and Applications”*, 24(1):85–90, 2017.
- [5] X. R. Cao, Basic Ideas for Event-Based Optimization of Markov Systems, *Discrete Event Dynamic Systems: Theory and Applications*, 15:169–197, 2005.
- [6] X. R. Cao, J. Zhang, Event-Based Optimization of Markov Systems, *IEEE Transactions on Automatic Control*, 53(4):1076–1082, 2008.
- [7] G. Lamperti, M. Zanella, X. Zhao, Diagnosis of Deep Discrete-Event Systems, *Journal of Artificial Intelligence Research*, 69:1473–1532, 2020.
- [8] D. Dobrev, Before We Can Find a Model, We Must Forget about Perfection, *Serdica Journal of Computing*, 15(2):85–128, 2021.
- [9] M. Schofield, M. Thielscher, General Game Playing with Imperfect Information, *Journal of Artificial Intelligence Research*, 66:901–935, 2019.
- [10] D. Dobrev, How does the AI understand what’s going on, *International Journal “Information Theories and Applications”*, 24(4):345–369, 2017.
- [11] D. Dobrev, AI – What is this, *PC Magazine – Bulgaria*, 11:12–13, 2000.
- [12] C. Wang, S. Joshi, R. Khardon, First Order Decision Diagrams for Relational MDPs, *Journal of Artificial Intelligence Research*, 31:431–472, 2008.