

Artificial Attachment: Will a robot ever pass Ainsworth's Strange Situation Test ?

Frédéric Kaplan
Sony Computer Science Laboratory Paris
6 rue Amyot 75005 Paris, France.
kaplan@csl.sony.fr

Abstract

This paper builds bridges between robotics, psychology and ethology in order to define a kind of "Turing test" that would constitute a possible aim for entertainment robotics. We argue that the design principles responsible for the success of existing artificial pet are not sufficient for building robots capable of a rewarding relationship with their owner. To achieve this goal, robots must display artificial attachment. Taking inspiration from the Ainsworth's Strange Situation Test used in psychology and ethology, we describe an experimental procedure to measure the attachment of a robot to its master. We then discuss how a robot can ever pass this test.

1 Introduction

What is the goal of an engineer that has to design an artificial friend ? This question is rarely asked. Recent years have been characterised by the expansion of animal-like, and recently child-like, robots. Some of this robots are presented as new educational technologies [7] but most of them are marketed as artificial companions [17, 9, 14].

If the purpose of an entertainment robot is primarily to create and maintain a relationship with its owner, how can we evaluate that a given robot is successful ? Artificial pets like Tamagotchis, Norns [12], or Furby Fables (see chapter "Furby Fables" in [23]) have been important commercial successes. Children and also adults to a large extent (see [17]), have started to spend a significant part of their leisure time engaging in a "relationship" with these

artificial creatures. Yet the richness of such a relation can be questioned.

The purpose of this paper is to propose a kind of "Turing test" [29] that would constitute a possible aim for entertainment robotics. We will first analyse the different design principles that can explain the success of existing animal-like robots and artificial pets. We will argue that despite their cleverness, these principles are not sufficient for building robots capable of a rewarding relationship with their owner. Analysing what is missing, we will argue in favour of robots capable of artificial attachment. Taking inspiration from a famous test used in psychology and ethology : the Ainsworth's Strange Situation Test [1], we will describe an experimental procedure to measure the attachment of a robot to its master. We will then discuss how a robot can ever pass this test.

2 Design principles of existing artificial pets

2.1 Uselessness and freedom

What makes artificial pets different from other robots or software agents is that they are not designed 'to be slaves'. This means that their first function is not to provide any kind of services such delivering mail or bringing coffee. In that sense, they are different from a lot of robots described in popular Science-Fiction. If AIBO is 'tired' of playing it won't play anymore irrespective of its owner's efforts to raise its interest. These creatures are not designed to respect Asimov's second law of robotics : *A robot must obey a human beings' orders* [2].

They are designed to have autonomous goals, to simulate autonomous feelings.

Engineers are not used to design apparently 'useless' creatures. This changes completely the way we should evaluate them. If you send an autonomous robot on Mars or on a serious mission in a nuclear plant, and the robot ever falls down, the engineer who designed it will feel he has not done his job well. It is not so serious when AIBO is falling. Usually people laugh. They are not expecting the robot to accomplish a very specific task, they just want it to be entertaining.

One way of showing that the pet is a free creature is to allow it to refuse the order of its owner. In our daily use of language, we tend to attribute intentions to devices that are not doing their job well. For instance, we do not develop any kind of relationship with our computer, washing machine or TV set when they work properly. It is only when they start disfunctioning, when they show that they can act differently that what we ordered, that we are ready to give them a kind of intentionality : "This computer refuses to work, the washing machine has decided to go on strike, etc."

The freedom of the pet, its apparent autonomy in the choice of its goals, seems a *necessary* feature for the development of an interesting relationship. Of course, it is not a *sufficient* feature. All the art of the creature designers rely on the way the pet will actually convince the user to interact with it.

2.2 Dependency

The paradox with an artificial pet is that although the creature is designed to be free, the owner has to have a reason for interacting with it. In most commercially available pets, repeated interactions are achieved because the owner feels responsible for his pet. The Tamagotchi for instance is a fragile being. If its owner does not give it the proper feeding, cleaning, nursing and playing, the pet will quickly die. A PostPet¹ might run away from its house, if its owner does not take care of it well. All these pets have been created to perform a kind of 'affective blackmail'. The owner must feel guilty if he doesn't take care of his pet.

¹Postpet is a mail software developed by Sony, in which a Pet is delivering your mail. It has been a important success in Japan. More information on: <http://www.sony.com.sg/postpet/postpet/index.html>

One very effective way of performing such a pressure on the user is to link the maturation of the creatures in some manner with the way the user is taking care of his pet. Most of the existing virtual or physical pets have a predefined maturation program which can be slowed down by a lack of interactions from the user. If you don't play enough with AIBO, it will not mature properly in the long run.

The trick is to create a positive feedback loop on the user investment in taking care of the pet. The more the user has spent time interacting with the pet the more it is crucial for him that the pet does not die or run away and matures properly. The initial investment may simply rely on the money spent to buy the pet. Then, the "relationship" emerges from this self-reinforcing dynamic.

2.3 Juvenile traits

In 1998, Bibendum, the character of the Michelin company, reached the venerable age of a hundred years. The same year, Mickey Mouse was celebrating its 70th birthday. In one article of its book, *Panda's thumb*, paleoethologist Stephen Jay Gould analysed the evolution of the physiomy of these two famous characters along the XXth century [11]. What is interesting is that instead of getting older, they seem to have become younger and younger. In the first films of the 30s, Mickey was a mouse with a long nose and small eyes. It was exuberant and sometimes cruel. Gould illustrates how year after year, as the mouse was pictured as a nicer fellow, its physiomy changed. In particular, its eyes and skull got bigger. A similar sort of metamorphosis happened to Bibendum, the Michelin's character. Initially it had a flat head and very small eyes. Now it is pictured with a round head and large eyes. It looks like a big baby.

For Konrad Lorenz, founder of ethology, such juvenile characteristics trigger innate responses in humans [21]. In other words, when we see a moving object that has similar traits to a baby's, we feel immediate and unconscious tenderness for it. Here are the characteristics that Lorenz mentions : "A relatively important head, an overdimensioned skull, large eyes placed low, small and thick extremities, a firm and elastic body, and inaccurate gestures".

These principles have been used for years by cartoon designers to make us feel sympathy for their

graphical creatures. They are now used by artificial creatures designers. It is important to note that these juvenile characteristics are independent from issues of realism. It is not because an artificial creature looks like a living being that it is going to be better accepted. On the contrary, it is possible that too much resemblances are bad for the acceptance of the creature. Furbies do not look like anything known. AIBOs do not have a fur. These are not random choices. By comparing the creature with its model, its limitations become obvious. Paradoxically, the creature has to be explicitly presented as artificial to be better accepted.

2.4 Emotional exchanges

There is a principle that artificial creature designers do not share with doll designers. Most dolls are characterised by their lack of expression. It is argued that this "neutrality" enables the child to project its own feelings and desires on the doll. As Serge Tisseron, a french mediologist, puts it "If the child talks to its stuffed toy and continues to believe it is listening to him, it is precisely because it never answers him" [27]

On the contrary, in artificial creature design, a creature with which a rich communication is possible is viewed as a potentially more interesting partner. Verbal communication is currently quite limited in most available artificial creatures. Some of them can produce preprogrammed sets of sentences (e.g. Furby). Some can understand prerecorded command sentences (e.g. AIBO 2). Current performances both in synthesis and in understanding are still disappointing. The main difficulty with verbal communication is that as soon as some sort of dialog is engaged between the creature and its owner, tolerance on errors becomes small. In particular, constraints on pragmatic relevance are very strong [6]. It is not acceptable that the creature should produce statements which are not relevant in the context of the conversation (e.g. saying spontaneously "the ball is red" in the absence of any particular motivation for talking about this subject). This explains why a lot of artificial assistants, which from time to time produce irrelevant answers or advises, are not well accepted by the users.

Because verbal dialog is too difficult, most artificial creatures designers have concentrated their effort into building creatures capable of interesting

non verbal communication. For instance, commercially available Aibo 2 implements a speech emotion recognition software and conveys its moods using body postures, simple melodies and light flashes. We can make at this point the same remark as the one we made concerning juvenile traits : Expressivity does not mean realism. Some cartoons characters can convey very precise and accurate emotions without using any understandable language. Some research prototypes have shown recently that it was possible to reach a satisfactory level of emotional interactions using rather simple techniques for emotion synthesis and recognition (see for instance the work of [5] and [22]). New products will certainly soon include similar techniques.

3 The design of attachment

Are these principles sufficient ? Is it enough that an artificial creature be built following these guidelines to gain our sympathy, to become a companion or a friend ? If we wish that the relationship between the robot and its owner approaches remotely the kind of social bonding that links a dog to its master, there is still a long way to go. But what is missing ?

3.1 What is attachment ?

Attachment is one of these ill-defined concepts, like intelligence or emotion. It was initially used to explain the bond that develops between a human infant and its caregiver [3]. The concept has been used a number of ways over the years. It is referred to a hypothetical factor that ties individuals together [20] or a behavior system that results in one individual seeking the proximity to another individual [4]. Instead of deciding on our own definition, we will try to characterize it through a set of examples.

Let's observe someone walking a dog. The dog walks sometimes in front of its master, sometimes behind. When it goes to explore bushes it always keep a look to check if its master is not too far away. This limit that the dog imposes to itself is the result of two contrary tendencies : its freedom and its attachment.

We appreciate when a "free" creature performs specific responses towards us only when it is not

forced to do so. Think of a squirrel that would come at your window, every morning. Although the "relationship" with the squirrel is rather limited, most people would appreciate such a daily rendez-vous, especially because the squirrel is not forced to attend it. It is freedom that gives value to attachment.

But imagine that the squirrel, maybe because you gave it some food, is now following you all day, at work, while doing shopping. It is then possible that you would less appreciate its continuous presence. The squirrel is not anymore a free creature that decides to spend time with you, it becomes something like a parasite, a servant or a slave. For this reason it is not interesting anymore.

Today's robots are either completely free (like most useless artificial creatures) or completely servant. No robot manages to maintain an interesting balance between freedom and attachment.

3.2 Can a robot be attached to its owner ?

How can a robot be well attached to its master ? For many people, this question is nonsense. A robot is a machine, an object. Some people may have a passion for objects. For instance some people that collect rare records, may well feel attached to some records of their collection. But even they are not crazy enough, to think, that the objects, in return, feel some kind of affection for them. Animals have a better status. The idea that they can give in return the love that we give to them is well accepted. For Dominique Lestel, a french ethologist, this possibility for reciprocity is fundamental to define our relationship with animals : "Interactions between men and animals becomes a state of connivence. Affective relationship overcomes intellectual relationship. For this reason, the animal is not an object nor a machine" [19].

But do we really know ? The debate on the possibility for a machine to feel attached recalls the one on its possible intelligence. In order to go beyond simple ideological or religious considerations, we would need a special "Turing test" for that purpose [29].

Such a test exists. It is called the Ainsworth's Strange Situation Test and it is used for checking, using rigorous criteria, whether the behavior patterns that a child displays in the presence and ab-

sence of his mother is normal [1]. The test consists of a succession of separations and reunions. The behavioral responses of the infant are classified into three overall patterns of behavioral organization :

- *secure* : The infant shows signs of missing after the separation, greets the parents when they are back and continues its normal activities (e.g. playing) afterwards.
- *insecure-avoidant* : The infant shows little distress at separation and avoids the parents upon reunion.
- *insecure-resistant* : the infant shows distress at separation and looks for contact in reunion but cannot be settled.

This test has recently been adapted to evaluate in the same way the attachment of a dog for its owner [28]. What we suggest is that it can also be used to test the attachment of a robot towards its owner.

3.3 Ainsworth's Strange Situation Test for robots

In trying to keep as close as possible to the basic experimental set up used in psychology [1] and ethology [28], we can define the following experimental procedure.

1. Introductory episode : The owner and the robot are introduced together to an experimental room. The owner puts the robot on the floor so that it can move freely.
2. Owner and robot : Initially the owner does not do anything particular while the robot explores. After 1,5 min a signal (a knock on the wall) is given to the owner to tell him to start playing with the robot
3. Stranger, owner and robot : A stranger enters and sits down. After 30s, the stranger initiates conversation with the owner. After 2 min, the stranger approaches the robot and stimulates playing. At the end of the episode, the owner leaves as unobtrusively as possible.
4. Stranger and robot / First separation : During the first minute, the stranger tries to engage

the robot in playing activities. If the robot does not play, the stranger can pet it. After 2 min, the stranger stops playing.

5. Owner and robot / First reunion : The owner approaches the closed door and calls the robot. The owner opens the door and pauses a moment to allow the robot to respond. The owner then greets and comforts the robot. Meanwhile, the stranger leaves. After 2 min, the owner leaves.
6. Robot alone / Second separation.
7. Stranger and robot / continuation of the second separation : The stranger enters. During the first minute, the stranger tries to play with the robot or to pet it. After the second minute, the stranger stops playing.
8. Owner and robot / Second reunion : The owner opens the door and pauses a moment before greeting the robot. Then the owner greets and comforts the robot. Meanwhile, the stranger leaves.

The exact duration of each episode still needs to be defined. This will depend on the speed of the robots. Experimental sessions are videotaped and the behavior of the robot is classified into behavioral categories such as *exploration*, *playing*, *greeting*, *passive behaviors* and *physical contact*. The relative percentage of time spent with these behaviors is established for each episode. By comparing these values during the different episode, it is possible to determine whether the robot is *secure*, *insecure-avoidant* or *insecure-resistant*.

With Ainsworth's Strange Situation Test we have a procedure for establishing in a rigorous manner whether a non human entity displays specific attachment patterns. Most robots would probably be classified today as *insecure-avoidant*. But if one robot shows similar behaviors than the one observed with secured infants and dogs, we would be forced to admit that from an external point of view, the robot seems to be attached to its owner.

4 The route towards attachment

How would it be possible to build robots that succeed in Ainsworth's Strange Situation Test ? We will not pretend to answer this question in such a short paper. We just want now to indicate possible milestones on the route towards attachment. Such a robot needs to recognize the presence or the absence of its master in an efficient way and react appropriately to it. To be able to do so, the robot will need to spend a lot of time with him. In consequence, we can identify three challenges. The first one is the organization of a self-reinforcing activity that involves many interactions between the robot and its owner. Typically, it would consist of a kind of ever evolving game. We will argue that one of the best examples of such a game is probably a training activity. The second challenge is the creation of an imprinting mechanism that enables the robot to learn to recognize the presence of its owner. The last one is finding the right balance in the robot's behavior between its autonomy and its attachment.

4.1 Training : an ever-evolving activity

Many a dog owner likes to teach tricks to their pet. In order to get a candy, the dog has to perform special routines or fetch certain objects.

At the Sony Computer Science Laboratory in Paris, we have designed several prototypes to illustrate how such a training is possible with an animal-like robot. For this matter we have collaborated with the ethology group of the Eötvös University in Hungary.

The current prototypes work on an enhanced version of AIBO, Sony's four-legged robot. We have decided to keep the original autonomous behavior of the robot and build our system on top of it. The system acts as a cognitive layer which interferes with the current autonomous behavior, without controlling it completely. This means that the trainer must take into account the global "mood" of the robot as it is generated by the motivation system. It is possible that a training session turns to be very inefficient if the robot is in a "lethargical" phase.

The first system that we have built is dedicated to object naming [15]. This system shares some similarities with the ones described in [25] and [8]. The owner of the robot presents a colorful object and says the word that should be associated with it. The robot analyses the image perceived by its camera and performs a simple segmentation using a *growing regions* algorithm. Because the robot's camera does not have a very wide view angle, only one object is generally in view. If several objects are segmented, one segment is chosen at random. The segment, which is supposed to be the topic of the interaction, is then analysed using a set of sensory channels corresponding to colour and shape properties. The corresponding sound is analysed by the speech system and associated with the simplified representation of the perceived scene. For instance, in the beginning the word "ball" might be associated with red shapes. As the robot plays similar language games, several views of the same object are associated to a single word like "ball". The set of all perceptions associated with the word "ball" defines an implicit category [13]. When a new object is perceived, the robot can try to recognize it using a *nearest neighbour* algorithm. The robot compares the segment seen with previous perceptions and utters the word associated with the ones which are closest to the one currently perceived. If it hears a congratulation afterwards it decides that it must be the right word, increases its "confidence" score, and will use it preferentially in the following interactions.

We have built another prototype dedicated for action naming [16]. This prototype is directly inspired from "clicker training", a method used efficiently by professional trainers for animals of different species [24, 26] (see also the model of [30] for virtual characters, also inspired by this training method). When a trainer wants to teach a dolphin to do a special jump on command, he cannot show it or explain it what to do. The constraints are very similar in our context. The robot needs to discover by itself what its owner wants. The idea of "clicker training" is to guide the animal using a signal meaning "go ahead you are in the right direction". Details about this particular training method go beyond the scope of this paper and can be found in [16].

These two systems, which are rather simple, do not pretend to solve most of the problems linked

with object recognition or action naming. Their purpose is to illustrate what kind of interesting interactions are possible between a robot and its owner. What is very important with these two models is that it is the *perceptual and social history* of the robots that determines its behavior in the future. As this history is unique for every robot, every robot will evolve differently. For this reason, taming robots is long-term but rewarding activity.

4.2 Imprinting

As the training sessions will pass, the attachment of the owner towards its robot should grow. But in the meantime, the robot will have a growing number of clues to identify its master in a satisfactory way (in particular his voice and his face). As it is learning to make the difference between the ball and its other toys, it will also learn to recognize its master.

In assuming that only sound is a reliable source of information for the robot, its task is to perform unsupervised speaker verification. The voice which it hears the most often during the training session will be assumed to be the voice of the master. Speaker verification has been the subject of active research for many years (see for instance [18] for a recent review of speaker verification techniques over the telephone). Yet, today no technique match well, to our knowledge, the requirements needed for robot imprinting which are to work (1) with speech captured in noisy conditions, and (2) that might be characterized by very different arousal and valence (typically comforting sentences or negative ones), (3) in an unsupervised manner and with (4) reasonable computing power. Following an approach similar to the one that Oudeyer follows for emotion recognition [22], we are currently doing an exploratory research in search for determinant features for robot-based speaker verification.

4.3 Balancing attachment and autonomy

Even if we assume that the robot is capable of recognizing its master in a satisfactory way, an important challenge remains : How will this capability affect its behavior ? To succeed in Ainsworth's Strange Situation Test a robot must neither be classified as insecure-avoidant (the robot acts as if its

- [11] S-J. Gould. *Panda s thumb : More reflexions in natural history*. W.W. Norton and Company, 1992.
- [12] S. Grand, D. Cliff, and A. Malhotra. Creatures: Artificial life autonomous software agents for home entertainment. In *Proceedings of the First International Conference on Autonomous Agents*,, pages 22–29, New York, 1997. ACM Press.
- [13] F. Kaplan. A new approach to class formation in multi-agent simulations of language evolution. In Y. Demazeau, editor, *Proceedings of the third international conference on multi-agent systems (ICMAS 98)*, pages 158–165, Los Alamitos, CA, 1998. IEEE Computer Society.
- [14] F. Kaplan. Free creatures: The role of uselessness in the design of artificial pets. In T. Christaller, G. Indiveri, and A. Poigne, editors, *Proceedings of the 1st Edutainment Robotics Workshop*. GMD-AiS, September 2000.
- [15] F. Kaplan. Talking aibo : First experimentation of verbal interactions with an autonomous four-legged robot. In A. Nijholt, D. Heylen, and K. Jokinen, editors, *Learning to Behave: Interacting agents CELE-TWENTE Workshop on Language Technology*, pages 57–63, October 2000.
- [16] F. Kaplan, P-Y. Oudeyer, E. Kubinyi, and A. Miklosi. Taming robots with clicker training : a solution for teaching complex behaviors. In M. Quoy, P. Gaussier, and J. L. Wyatt, editors, *Proceedings of the 9th European workshop on learning robots*, LNAI. Springer, 2001.
- [17] M. Kusahara. The art of creating subjective reality: an analysis of japanese digital pets. In Maley C. and E. Boudreau, editors, *Artificial life VII Workshop Proceedings*, pages 141–144, 2000.
- [18] L.F. Lamel and J.L. Gauvain. Speaker verification over the telephone. *Speech Communication*, 31:141–154, 2000.
- [19] D. Lestel. Des animaux-machines aux machines animales. In B. Cyrulnik, editor, *Si les lions pouvaient parler : essais sur la condition animale*, Quarto, pages 680–699. Gallimard, 1998.
- [20] K. Lorenz. *On aggression*. Harcourt, Brace and World, Inc., New York, 1966.
- [21] K. Lorenz. *Studies in animal and human behaviour*. Harvard University Press, 1970.
- [22] P-Y. Oudeyer. Emotional interactions with humanoids using speech. In *Proceedings of Humanoids 2001*, 2001.
- [23] Mark Pesce. *The playful world : How technology is transforming our imagination*. Ballantine books, 2000.
- [24] K. Pryor. *Clicker training for dogs*. Sunshine books, Inc., Waltham, MA., 1999.
- [25] D. Roy. *Learning form sights and sounds : a computational model*. PhD thesis, MIT Media Laboratory, 1999.
- [26] Peggy Tillman. *Clicking with your dog*. Sunshine Books, Walthman, MA, 2000.
- [27] S. Tisseron. *Petites mythologies d aujourd hui*. Aubier, 2000.
- [28] J. Topal, A. Miklosi, V. Csanyi, and A. Doka. Attachment behavior in dogs : a new application of ainsworth’s strange situation test. *Journal of Comparative Psychology*, 112:219–229, 1998.
- [29] A. Turing. Computing machinery and intelligence. *Mind*, 59:433–460, 1950.
- [30] S-Y. Yoon, R. Burke, B. Blumberg, and G. Schneider. Interactive training for synthetic characters. In *AAAI /IAAI 2000*, pages 249–254, 2000.