Phase transitions

A Forrest Gump-like account on AI 1984-2024

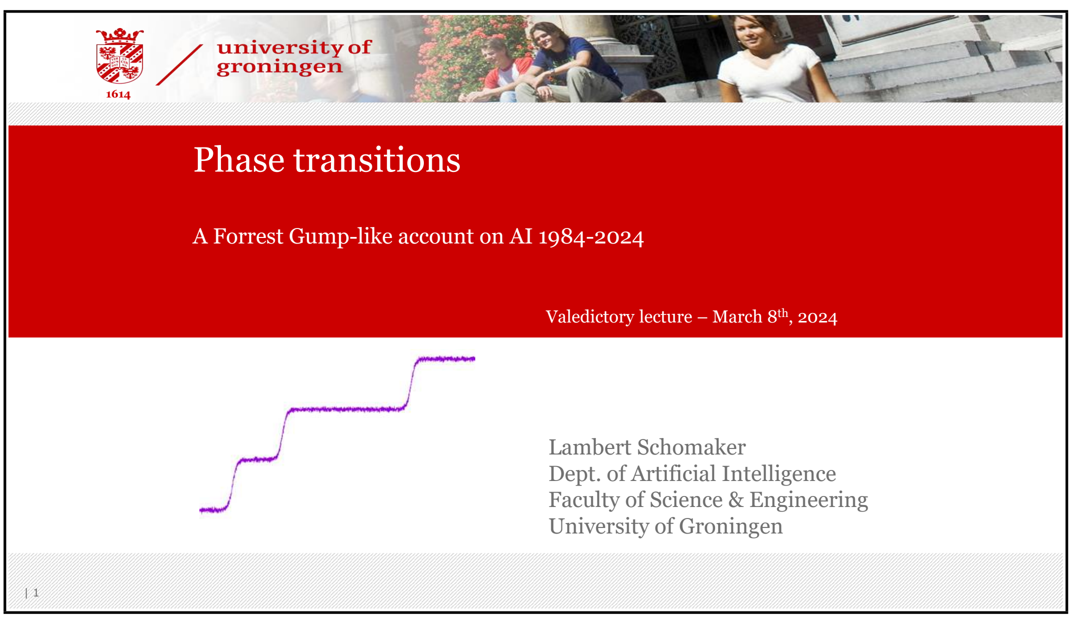
valedictory lecture

Lambert Schomaker

March 8th, -2024

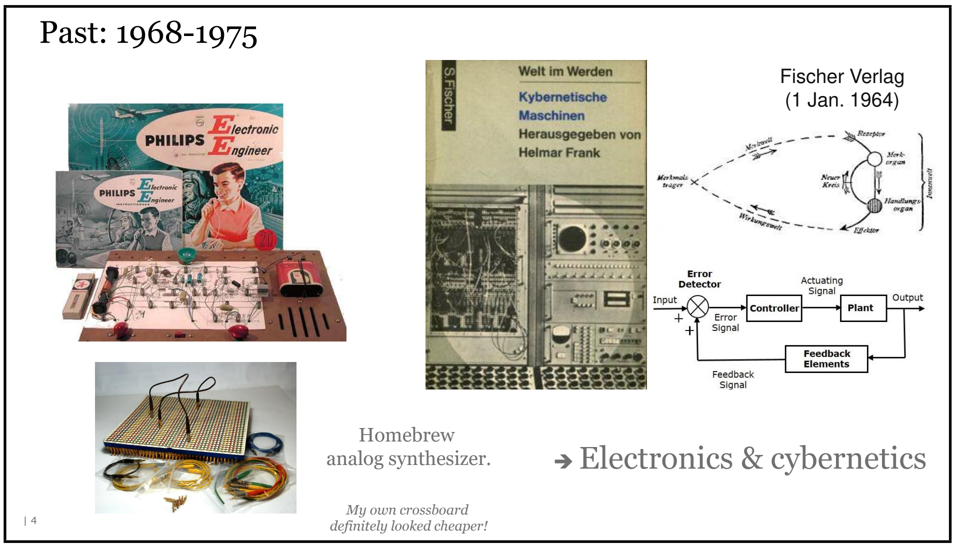
Preamble

Welcome to this speech which I have baptized as "phase transitions", a theme that will reoccur during this lecture and I also call it the "Forrest Gump" account on AI because, starting in 1984, now we have 2024, there are a lot of things that happened in AI and on several occasions I was also there, like the movie character Forrest Gump. First, we have my early life, that's just the ordinary early life of a village boy with a lot of freedom to go into nature which - as has become clear - I can now do in Sweden again. After the initial youth, there is an aspect of "electronics" and "cybernetics" and the next slide makes that clear.



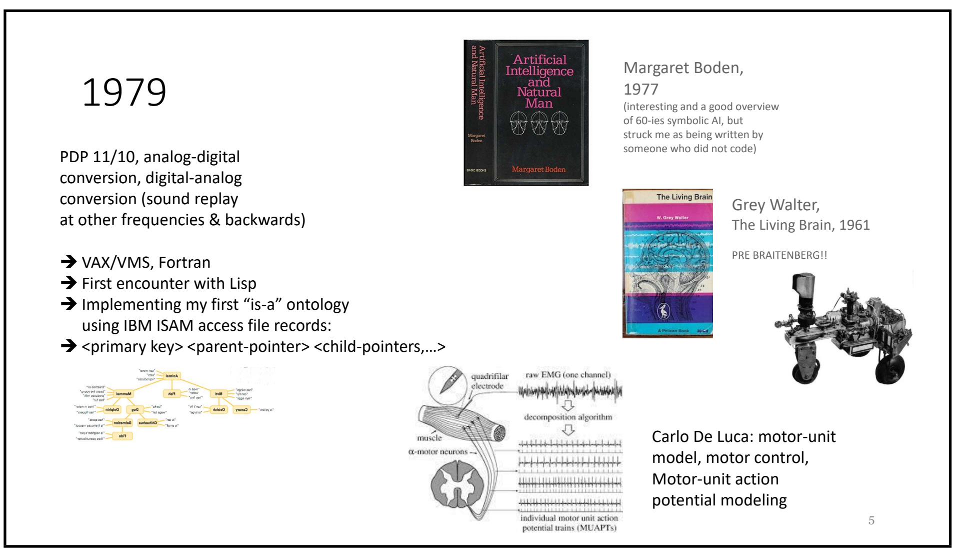
slide #4

So I got this Philips Electronics experimentation box that was very important to me, and then an uncle gave me this German book "Kybernetische Machinen", my uncle in Germany was an engineer himself and there I learned about feedback. Feedback and electronics together is already quite interesting for control of systems with for instance with light dependent resistors (LDRs) - I am not going into the details - and then of course at High School what we all wanted to have is a Moog synthesizer which was too expensive, so I tried to make one with a crossboard. I didn't have this kind of luxury crossboard (photo on slide) myself but it's all a question of connecting inputs and outputs with interesting analog elements in the middle. This slide, I'm not pondering too long on it, is important because it shows really the roots of where I started from.

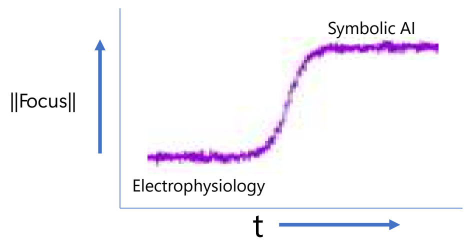


slide #5

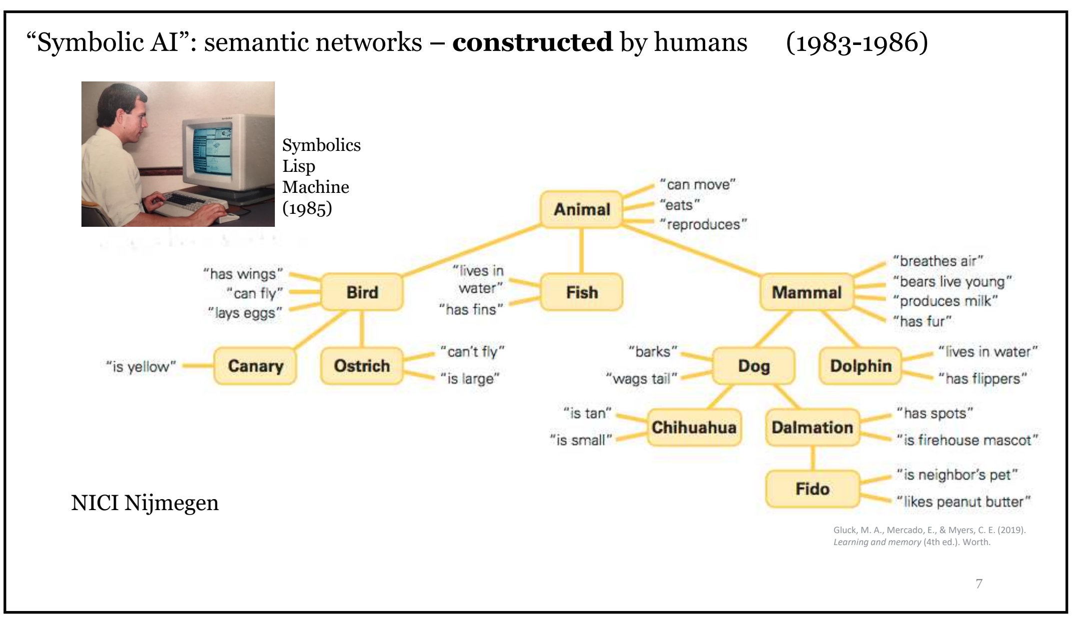
By 1979 I was already at University, I was doubting whether I should leave psychology and go to the Technical University of Eindhoven. Then the physiology department head, Professor Cees Bruna bought this kind of multimillion Guilder computer that was even faster than the one at the Computing Center, i.e., a VAX computer. So I learned to program in Fortran. Lisp the language of AI was also a first encounter. Then I tried to make this kind of knowledge and semantic graphs in using a specific technology, IBM indexed sequential files - going into the detail but the nerds among you will know what I'm I'm saying here. At the same time I read the book of Margaret Boden on AI because by that time there was already already a history of AI in the 1960s. I was struck, because I believed that programming was everything so I found it quite peculiar that she wrote this book without knowing the details of coding but of course it was a very inspiring book. This is on, say on the symbolic side, but we also have Grey Walter's "The Living Brain" with this kind of simple turtle robots, pre-Braitenberg (1984). Also - I'm not going going too much into the depth but - it strikes me that always, well often, the original paper and the original work is forgotten somehow. In neural networks this is also happening. Then my real work is in a corner here on the slide but that was my real research, which is modeling the spiking behavior in the motor system and what was - I found - really very exciting was that there was a mathematical model by Carlo de Luca on the motor units and neurons and, lo and behold, the model also really predicted what the neural signals would look like. This was very exciting but now at the end of that stage I needed to look for a job and of course there was no job in electrophysiology but there was a job in cognitive science in Nijmegen.



slide #7

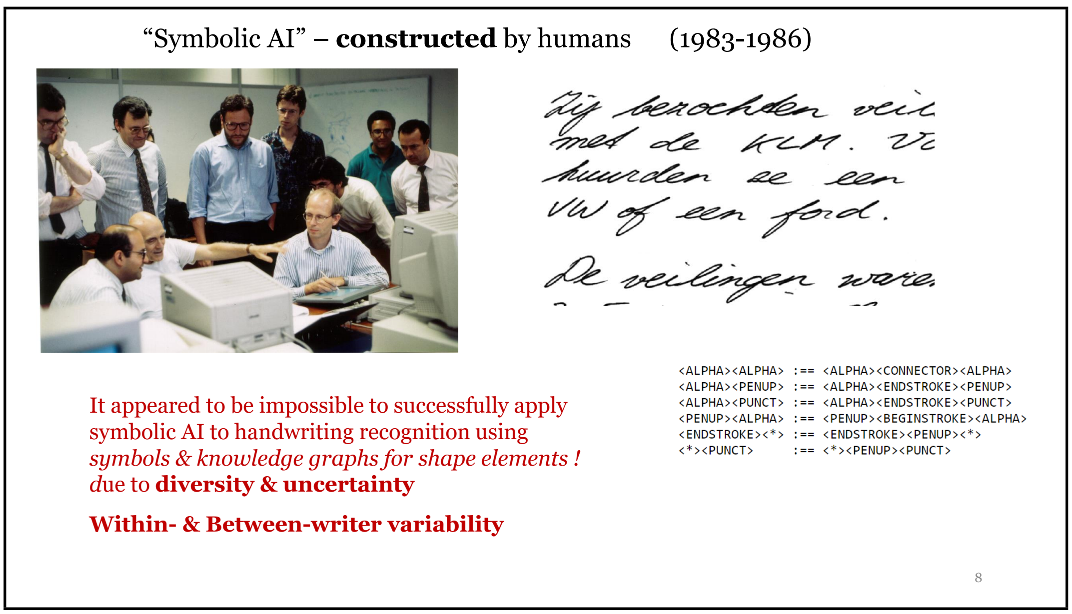


And so there was a shift, a phase transition from electrophysiology to symbolic AI and in symbolic AI what you do is you have of course a Symbolics Lisp machine machine in that age (photo on slide). So - for the nerds again - the machine language in this machine is Lisp: Everything is Lisp from the bottom up and then you can type in knowledge graphs and as you can see in symbolic AI (points at slide) you have Abstract Concepts here like Animal, and Animal can be subdivided into subclasses Bird, Fish and Mammal and within those subclasses we have subclasses again and then in the end we have the instances like Fido the Dalmatian and so on. But there's an important sentence here it says "Constructed by humans"



slide #8

And "Constructed by humans" is interesting also if you talk about handwriting and the belief was, even within the field of pattern recognition, that you should do everything explainable, explicit and symbolic rule-based, with knowledge graphs like the ones that I showed and so I tried in a European project to do things in handwriting recognition with tablet computers. You see here the people standing around me. This is me, for those of you who don't see the resemblance and I really I really did my best to make a semantic model, or a syntactic model or in any case a symbolic model of of handwriting and it utterly failed. You can do something in a very clean handwriting style without noise etc., but if a new writer comes in you need "Construction by humans" again, because each and everyone of you is doing very detailed, different things in your handwriting and in this symbolic paradigm the only thing you can do is to hand code it yourself and of course that's really a problem. So "diversity" and "uncertainty" are really a problem and and not only between writers but even if one person writes several times, then there are also many, many differences.

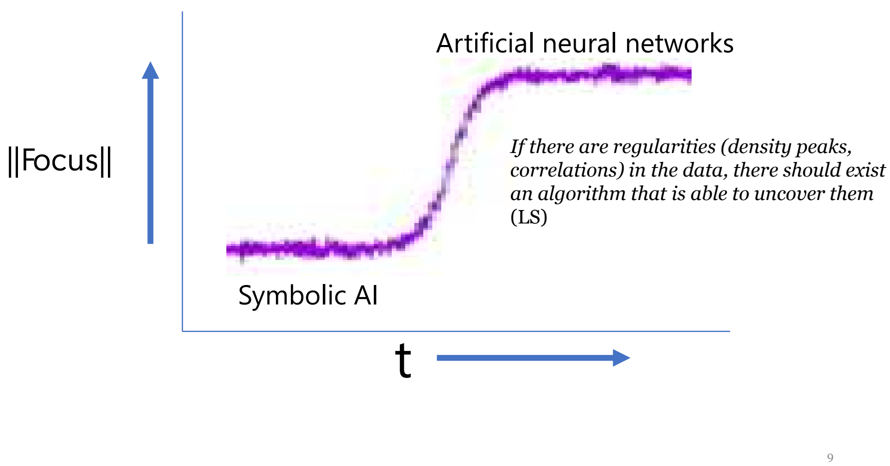


slide #9

Therefore we go into the second phase transition from symbolic AI to artificial neural networks and there's a kind of uh it's not a theorem or an axiom but it's a kind of informal claim that I formulated, also to convince people:

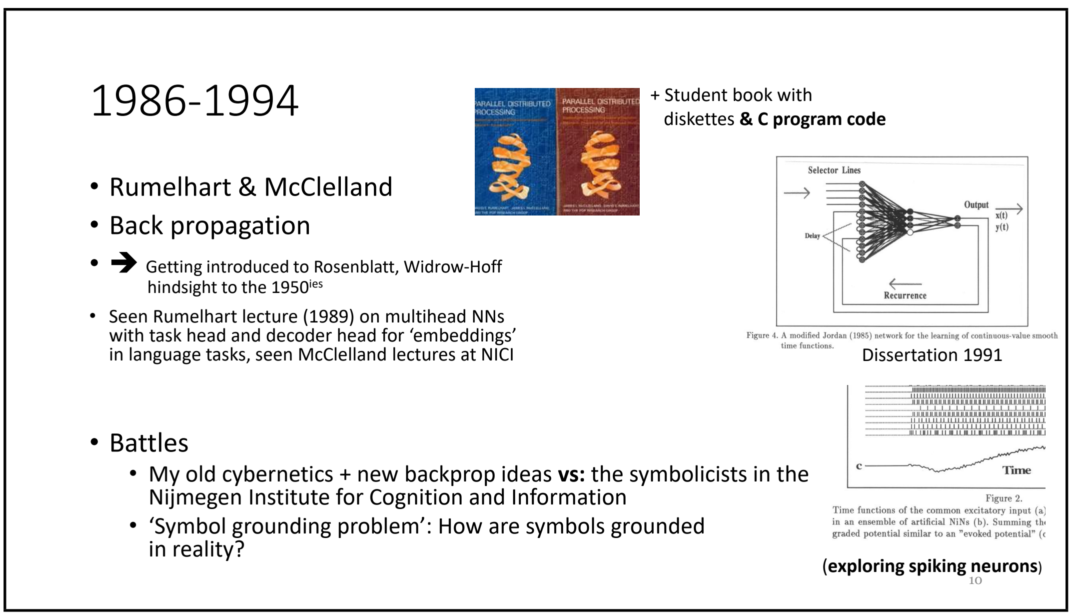
> "If there are regularities (and regularities are density peaks in distributions or correlations between dimensions) in the data, there should exist an algorithm that is able to uncover them".

Of course it has to do with the word regularity. Traditionally, this was not liked that much. The word that people in engineering these days use is "data driven" but of course we call it different things in machine learning, but indeed the data should tell what is the structure in there and not a kind of normative tree that maybe looks nice but doesn't cover the variation in nature.



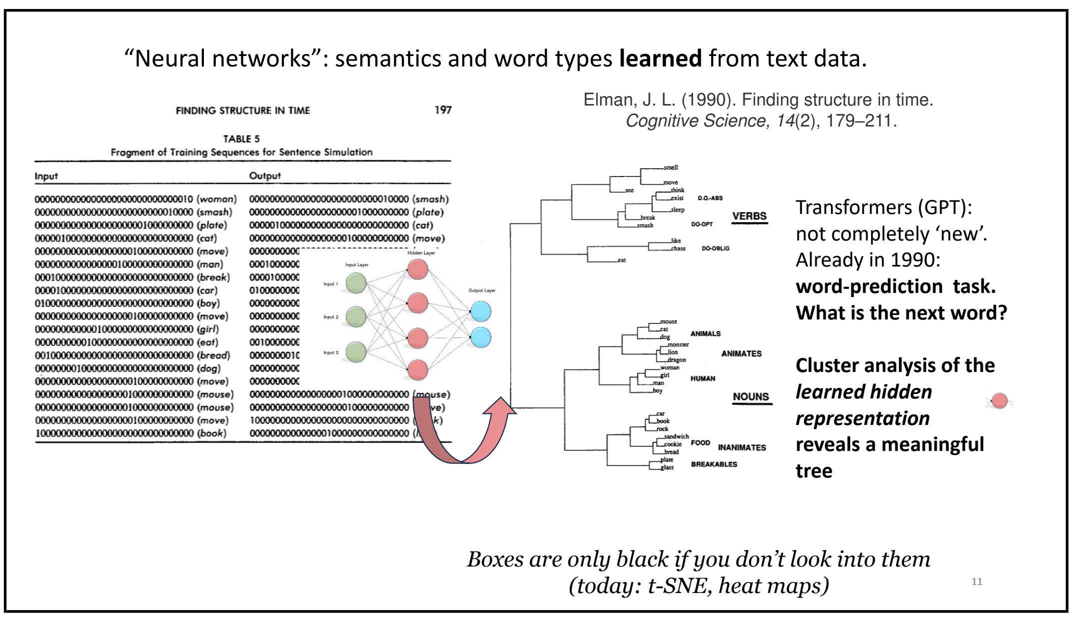
slide #10

In 1986, when I was at a conference in Montreal my colleague Professor Pietro Morasso from the University of Genoa, from a robotics Institute I also collaborated with, he dropped two big Bibles (pointing at Rumelhart & McClelland books) on my desk and there was also a student book with C program code so this is heaven, right? Because it's not only yaddah yaddah about the theory you can also test it, and so in my dissertation I started to experiment with recurrent neural networks, but also because of my electrophysiological past I did modeling with spiking neural networks so I learned a lot, you learned about back propagation, you looked back into the 50s, like Grey Walter whom I already mentioned but also Rosenblatt and Widrow-Hoff, other famous people in neural networks that that I learned to know. And this is kind of of a Forrest Gump like event so I really had these lectures by Rumelhart and by McClelland and just also for the younger people, already in 1989 multihead neural networks were being used with the multitask setup for an autoencoder and a classification task simultaneously. So a lot that is now sold as completely new, with new names attached to it in the literature, a lot of that you can just find in those old books. It's already there but it's reinvented, gets slightly different names, and uhm, this is true in several cases. Now the advent of these systems that are learning structure from data was of course met with a lot of internal battles. There was a battle between, let's say, my generation, and the ones who had learned from Minsky and Papert (1969), i.e., that neural networks were rubbish but also the psychologists who did like neither the symbolic modeling nor the neural networks. So it was a, well, interesting time so to speak.



slide #11

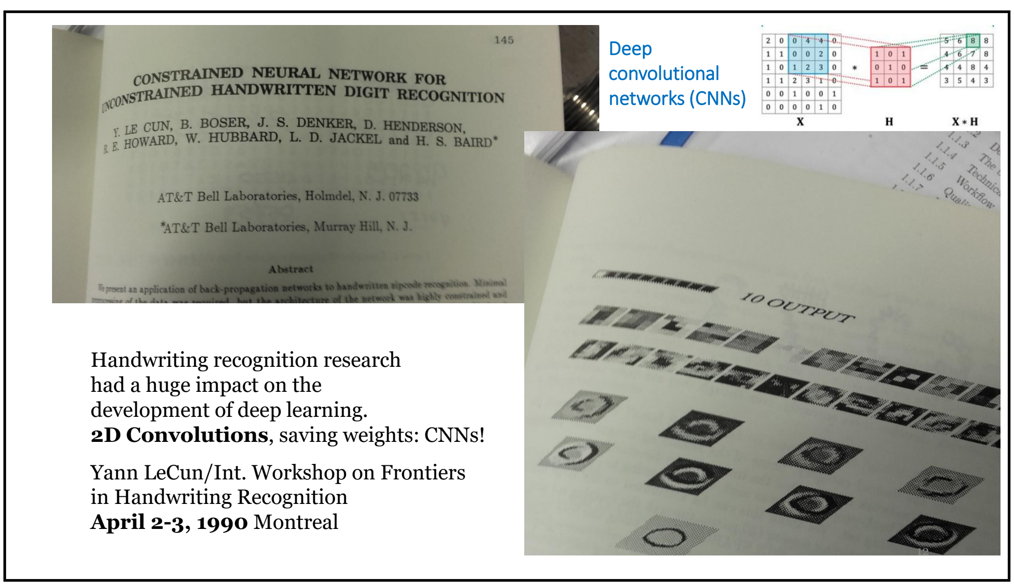
And now, to tell even a stronger example of things already in that gray past being very similar to today, is this seminal paper that I can really advise to everyone to read again: The Elman 1990 paper "Finding Structure in Time". So what he did is exactly the same as what the large language modeling people are doing these days. Namely, you have sequences of words, e.g., "woman smash plate", "cat move man", "break car boy" and so on, and what what happens is that the neural network needs to predict the next word. Now we get to the magic. Because if you train this task to the system and it has inputs which are these patterns, it predicts the next time step. If after training you analyze with a hierarchical clustering technique what the neural network has learned (pointing to the hidden layer). Then, without human construction a tree pops out where the system has itself learned that there are "verbs" and here are "nouns" and that within the nouns we have the "animates" and the "inanimates" and so on. So, yeah I really didn't understand why people were **not** enthusiastic about this, absolutely. And the insight is, of course that with neural networks the common criticism is they are black boxes, you don't know what they're doing. Yes, if you don't look into all those matrices and vectors, then you don't know what they're doing! But if you open it up and you do normal statistical and signal analysis of what is happening in the inside, you will learn a lot. But neural networks are complicated and it's difficult to understand everything. But this whole idea that it is a black box and you don't know anything, that's not true. Maybe you there are no black boxes, maybe there are lazy scientists, right, who don't do this analysis!



slide 11

slide #12

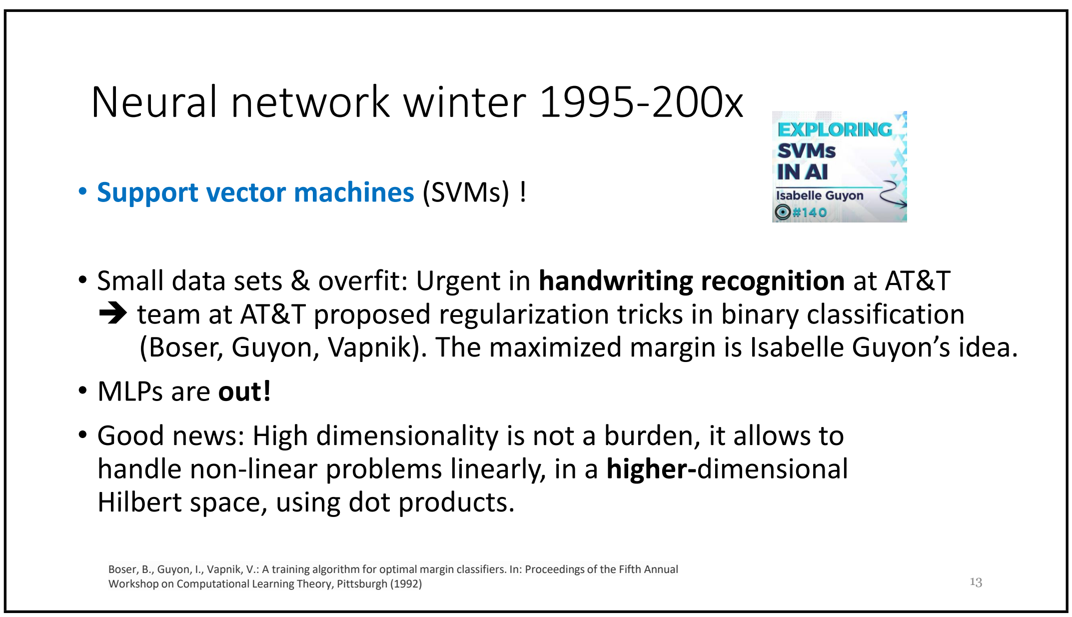
Handwriting recognition studies were shown today (in the presentations this morning). I'm quite happy with that, because handwriting recognition on the one hand has a kind of dusty image, old-fashioned stuff, who does it any more? Maybe it has to do with archives etc... But, because handwriting recognition is so terribly difficult, there are many, and more than I'm mentioning here, many innovations in machine learning which are from the field of handwriting recognition. So this is from a presentation by Yann LeCun (1990), he came up this morning in the PhD presentations. He introduced the convolutional neural network already in 1989. I myself I heard first about it - this is the Forrest Gump again - at a conference in Montreal in 1990 and here you can quickly and easily see it: You see the title, this is the original title, and you see here the receptive patches in the hidden layers, going towards the the abstract classes and then here a linear strip at the output. It is the digit recognition task that everyone knows by now, the MNIST data set with the digits 0 to 9 and so this innovation is - I'm not going into the detail for everybody - that before, for this classification task from this pixel matrix to this matrix in the next layer, all units, everything would need to be connected with everything. So from the coefficient-estimation side this is undoable because quickly you have thousands of coefficients or weights which would mean that you would need tens of thousands or hundreds of thousands of samples to have reasonable results. What does Yann say: I don't need to propagate the whole image I can just propagate small patches of the image. I have here a kernel of only three by three weights and as long as these kernels detect meaningful structural pieces of information in the data and propagate it to the output I can do the task: I don't need to connect everything with everything. There is some kind of precursor work by (Kunihiko) Fukushima: The neocognitron has similar properties. There this is kind of seminal work. But there are people in Germany who are against this rendition of history, notably Jurgen Schmidhuber and he is so adamant about it that there is a new verb that is now invented it's called Schmidhubering and that means that you are pushing a point that the history of science went different than it really did.



slide 12

slide #13

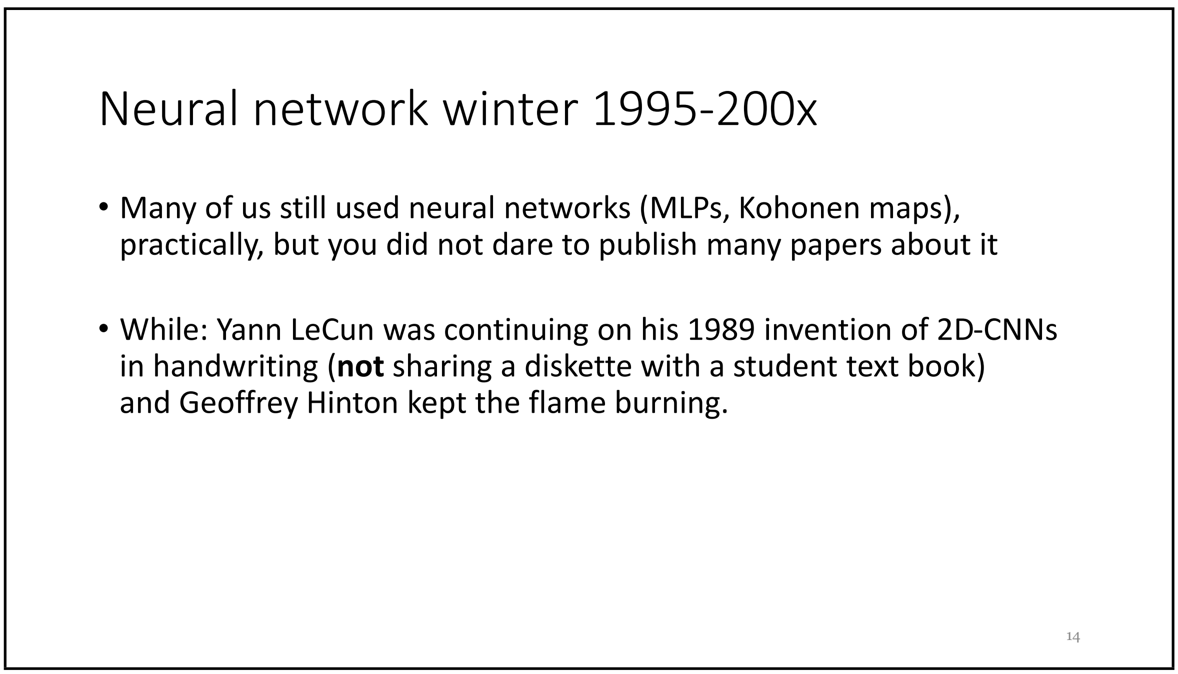
Now, of course, it was not completely rosy because computers were not fast and the data sets were limited and also very unrealistic: They're too too clean and what happened at AT&T, where they had a handwriting recognition group because - of course the Apple iPad is now also almost history but everyone wanted to have a tablet on which you could make gestures or write. So everyone was trying to do that and the product managers at AT&T they were very upset by the fact that if you train a neural network twice you may have a different result. And in standard engineering and in mathematics, I see Henk Broer nodding, that this is very undesirable right? So you you don't want to have that and so they called in a group with Russian mathematician Vladimir Vapnik and they said: "Can you not make a kind of classifier that doesn't have this abundance of coefficients and is more stable in the training. This has become the Support Vector Machine by the team of Vladimir Vapnik, Bernard Boser and Isabelle Guyon. The maximized margin is Isabelle Guyon's idea, which I specifically say because this is also kind of a Rosalind Franklin-type of event where there is a famous paper and an important component of the invention is the maximized margin - I'm not going into detail now - but this is her invention. The consequence is that the multi-layer perceptrons, i.e., the normal neural networks are out at that point. But still, for us in machine learning there's a kind of indirect good news because the support vector machine worked not by reducing the dimensionality of a problem but by enlarging it. The trick made it possible to separate classes of shapes in a high dimensional space with a simple hyperplane instead of a very curved manifold between the classes using Hilbert space or dot products. It will become too technical but this is a kind of nice result of that stage.



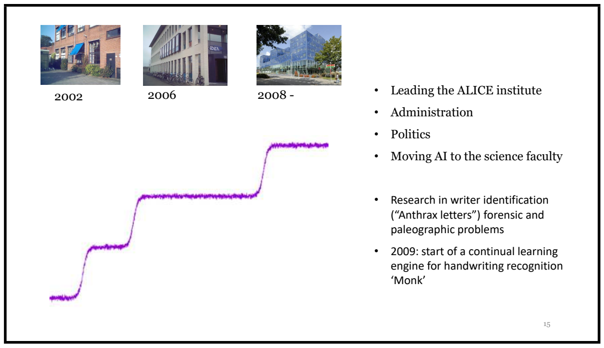
slide 13

slide #14

But many of us - we did use the neural networks, not only the multilayer perceptrons but also Kohonen maps that were mentioned today - but many, I myself also, I didn't dare to really publish about multilayer perceptrons, it was like kind of old fashioned already and it didn't, uh, the asymptotes of performance were reached, and there was not a lot of movement. And Yann LeCun - the guy who's now working at Meta as the director of research - he was just continuing it. He did something different in that he didn't have a book with a diskette, so in the decade 1990 to 2000 all of us were trying to reverse engineer what Yann LeCun was doing but like Marius Bulacu also said it was quite difficult to get it really done. Now today you have GitHub, now everything is popularized but it was secretive also because of course Yann was working for AT&T so not everything was open. Also Geoffrey Hinton kept going, who needs to be mentioned. He was was also one of the authors in these two neural network bibles that I just showed.

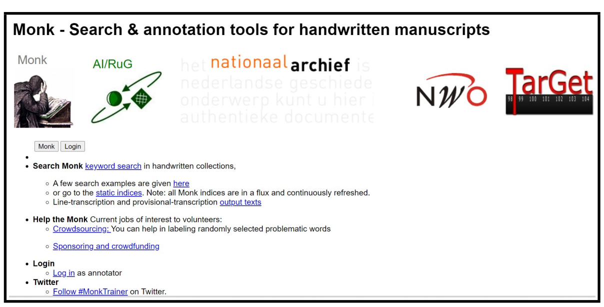


slide #15



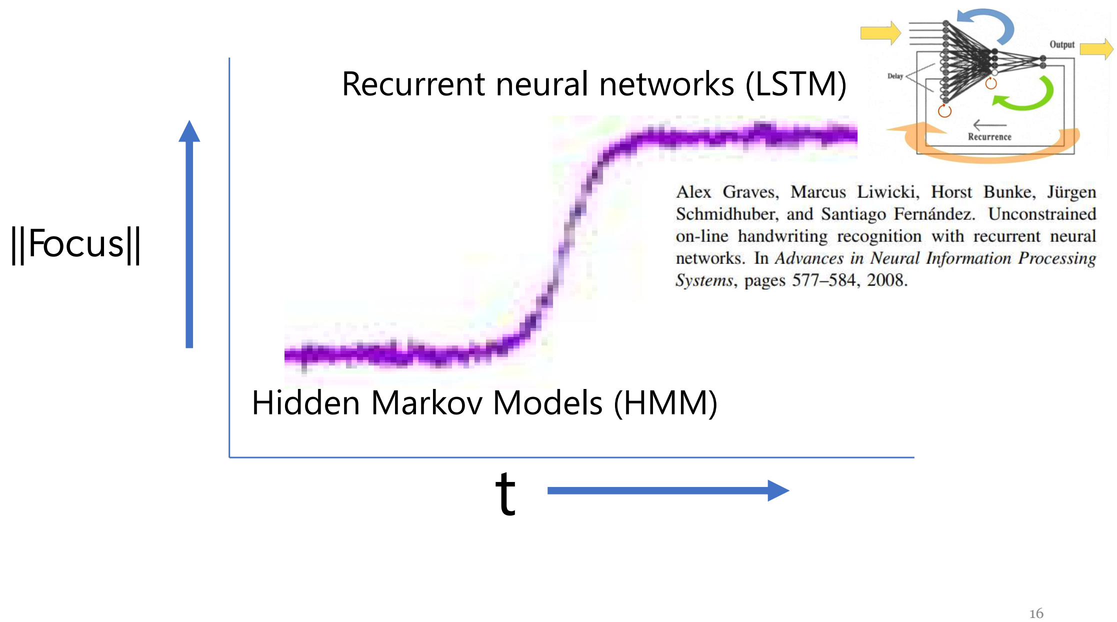
slide 15

Very briefly I show you the phase transitions at the administrative and 'leading the Institute' level, a lot of politics, moving AI to the science faculty which was quite a complicated thing. I still kept doing science in those days but of course as Niels Taatgen already said, the rate at which you can do it - he knows it by now himself as well - there is a limit to what you can do in research. Also I'm happy, we didn't agree on this in advance, but Marius mentioned the "Anthrax letters". I tried to shed off the handwriting-dusty thing and go to computer vision and robotics but, somehow it always came back to me and we continued research in that area. In 2009 I started the kind of continual learning engine for handwriting because my idea was that well machine learning and deep learning works but I need a kind of harvesting method for labeled data. Using nearest-centroid matching and presenting hit lists of image on a web site, volunteers can easily label handwritten word images. Then neural networks and other techniques can be used. This system is called "Monk" (Schomaker, 2009, 2016) and it was made possible by a project that I had also with professor Edwin Valentijn, also in the room, in the "Target" project with the astronomers. Maybe you remember, Edwin, that you said mmmh, this is interesting we can use this this to show that we're not only looking at the black holes and galaxies in this grant proposal but we also have a cultural heritage connection: "We allow you in the noise" in terms of the size of the data, i.e., the disk size needed for handwriting in comparison to all the astronomy images. I'm very grateful for that because it had a lot of effects.



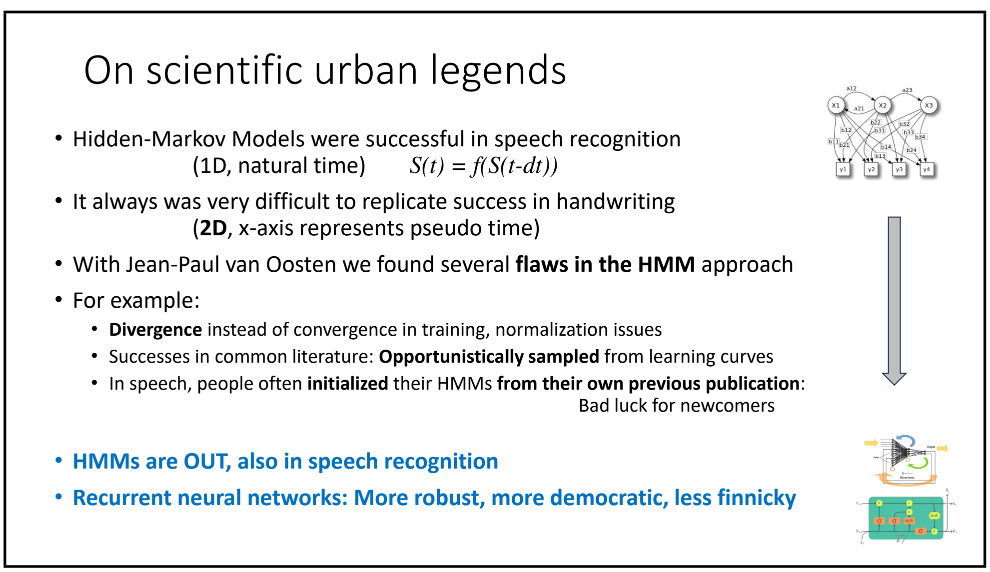
slide #16

Now within the AI field, there were other things going on. This is another another important phase transition that's from the Hidden-Markov Models to recurrent neural networks. So in 2008, there's a paper by Alex Graves, Marcus Liwicki, Horst Bunke and Juergen Schmidhuber - there there he is again, he is a good researcher right - And the idea is that you have a stream of data and that you feed back, over time steps, previous information into the network and process it again. I also see Arjan van der Schaft in the room, so you know that recurrence also implies differential equations and instability and and problems but the point is that these techniques were much better than the hidden Markov models.



slide #17

Now this is work where I want to talk about urban legends. With scientific urban legend I mean a kind of concept that may be rooted in mathematics and then everyone believes that you have to do it this way because there are proofs like the Baum & Welch convergence proof. In speech recognition people did it and it was relatively successful. Basically the model is that the current state $S(t)$ can be computed from just the previous state $S(t-dt)$. But speech is in natural time and when you try to apply this to handwriting images there was no low-hanging fruit, there was no easy success. Handwriting is two-dimensional. If you want to come up with a time axis it will be a pseudo time axis like the x-axis. And it was extremely difficult. There were some people in our field who had success with Hidden-Markov Models but I really didn't understand why everyone was running after Markov models. It was solid - who would want to argue with Bayes' Theorem right? But then we had a project with Jean-Paul van Oosten. He is already also here, and we just found several serious flaws in the Hidden-Markov approach. For instance, what Jean-Paul showed is that in the training process there is **divergence** instead of the promised convergence by Baum and Welch; there are normalization issues; then we saw a kind of ethical things in scientific practice. In neural network training the loss curves have a little bit of variation but they just go down and you have an asymptote and then the loss is minimized. In Hidden Markov the comparable 'loss' curves are very irregular. So what happened in literature is that people just look at that curve, then select some kind of model that looks good and then report on it. That's the second kind of fundamental problem and then the worst of all is, that we found out that the speech recognition researchers often initialized their hidden Markov models from the best result in their own previous publication. So you don't start from scratch, randomly as as you are supposed to do, but you inherit the good results. This means: "bad luck for newcomers"! Is that really scientific? So I find this very intriguing. On the one hand this claim of scientific, mathematical rigor and on the other hand something that is the practice, which is completely different. So: Hidden Markov models are out, also in speech recognition. These recurrent neural networks, any student can pick this code and try something. It is more robust, it is more democratic and less finicky.

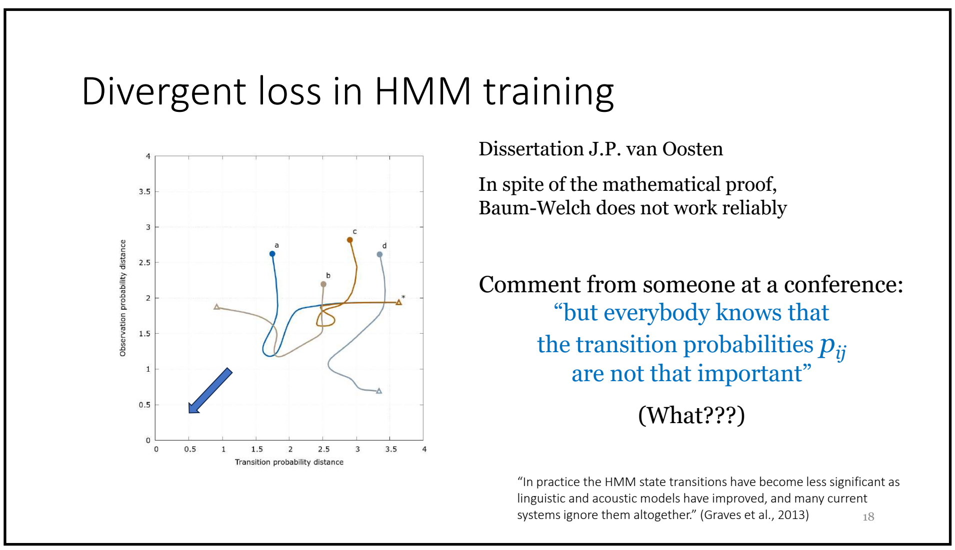


slide #18

Just to give an example of what Jean Paul found out: (points at slide) these are trajectories of relaxation so here on the y-axis are the observation probability differences that you need to achieve, they need to be low, and this is the transition probability distance on the x-axis. An ideal model would migrate the solution to the lower left here, to zero. But we can see all kinds of trajectories, which contrary to Baum & Welch and their proof goes in all kinds of directions. Now you also immediately see what the speech-recognition people did because they did a lot of experiments here and then made sure for the next paper that the seed point is closer to the lower-left corner and then of course you have better results. Now this is too much detail, the next slides will be faster. But we went to a conference and maybe also Jean Paul remembers things about that, we were at a conference and then someone said suddenly: "But everybody knows that transition probabilities P_{ij} are not that important". WHAT? The Markov model is about time. The transition matrix with elements P_{ij} is the **only** aspect of the model that really handles with the sequence of time, the other probabilities have to do with observation probabilities and per-hidden state conditional distributions, things that are occurring but decoupled from time. So how is it possible that you have you have this kind of almost religious belief in the Hidden Markov approach with its ethical problems, convergence problems, and then in the end some kind of admitting "well, P_{ij} is not that important" I mean, we were really shocked. Here is a quote from an actual paper:

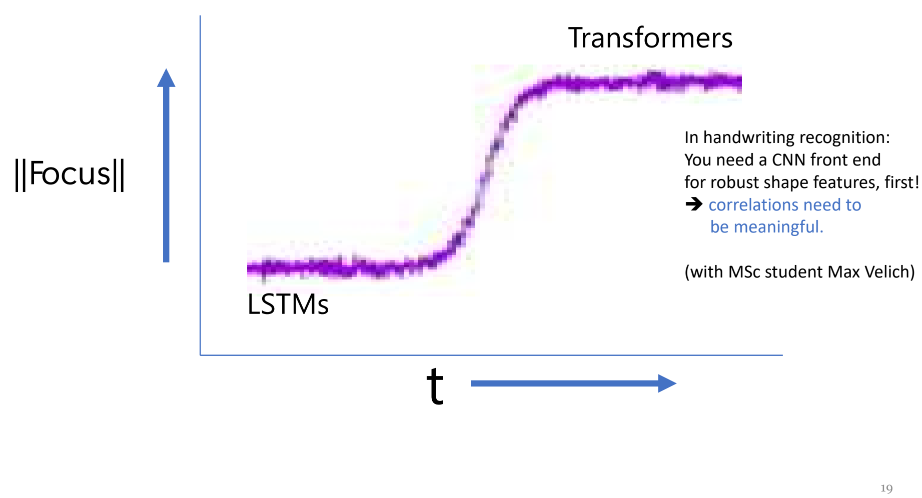
> "In practice the HMM State transitions have become less significant as linguistic and acoustic models have improved and many current systems ignore them all together."

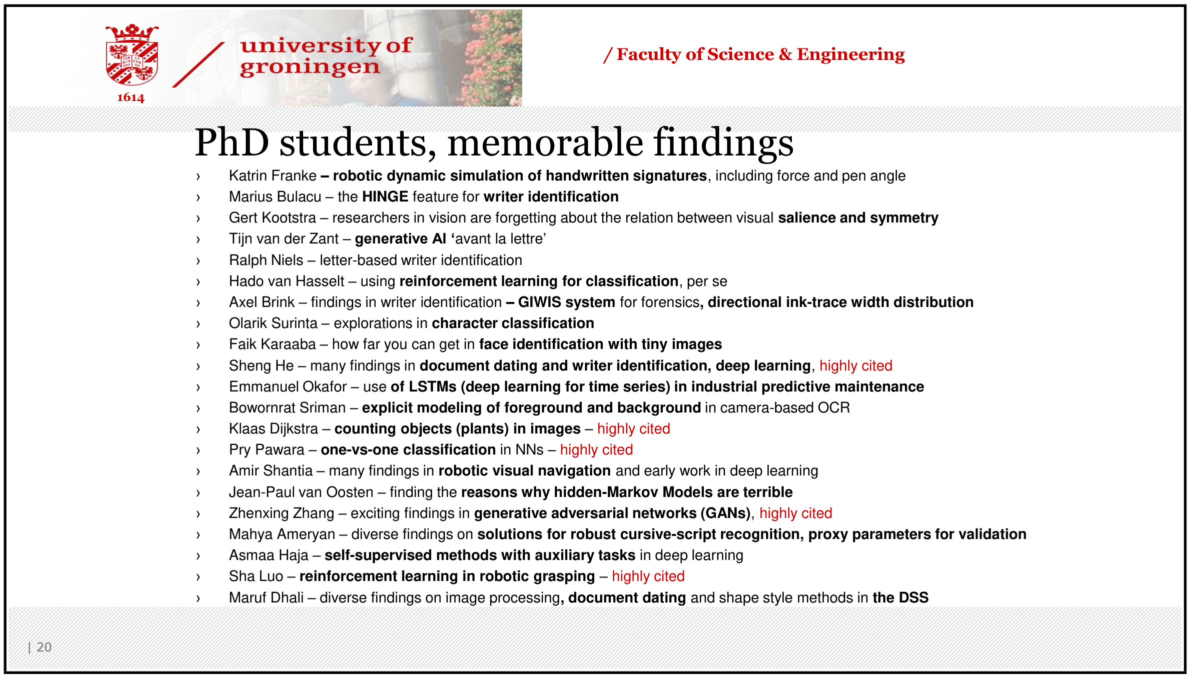
So so for years we wanted to be obedient and do things in handwriting recognition like the big successes in speech and in the end the neural networks won.



slide #19

Of course this is only just a moment in time because LSTMs, the recurrent neural networks, also have their disadvantages and they're now being replaced by Transformers of which I can say that in handwriting recognition mostly you need to refine the information that you have. You refine it, e.g., with a convolutional neural network to have a feature embedding with meaningful information along the dimensions because a correlation (a dot product) between two noisy variables is bound to be meaningless. Transformers are very nice, we all know GPT but also there, low-hanging fruit is not that easy in handwriting recognition. My student Max Velich had very good results with transformers.



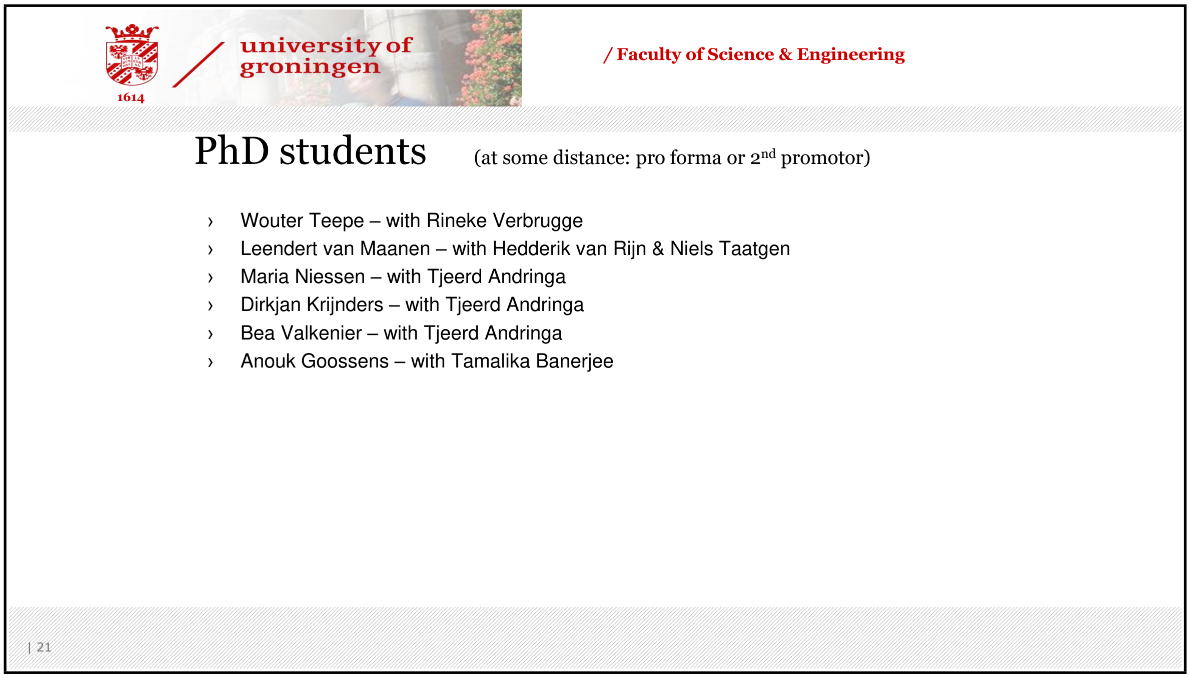


slide #20 - The PhD students

So these are not all my PhD students and many of them with memorable findings. Several of you were already 00:36:03,445 –¿ 00:36:06,765 mentioned. Hado van Hasselt ended at DeepMind and Sheng He, whom you saw in his video ended at Harvard Medical School. We have some others: Zhenxing Zhang - I'm going to present something on his work in a later slide. Mahya Ameryan is also present here. With her we found proxy parameters for validation that do not use labeled or supervised data, that worked quite well. Asmaa Haja was on video and Sha Luo as well. And Maruf Dhali, you have seen him several times today. Each of these PhD students - I cannot handle all of them in detail, I'm sorry guys! - had interesting findings. I made a selection to come.

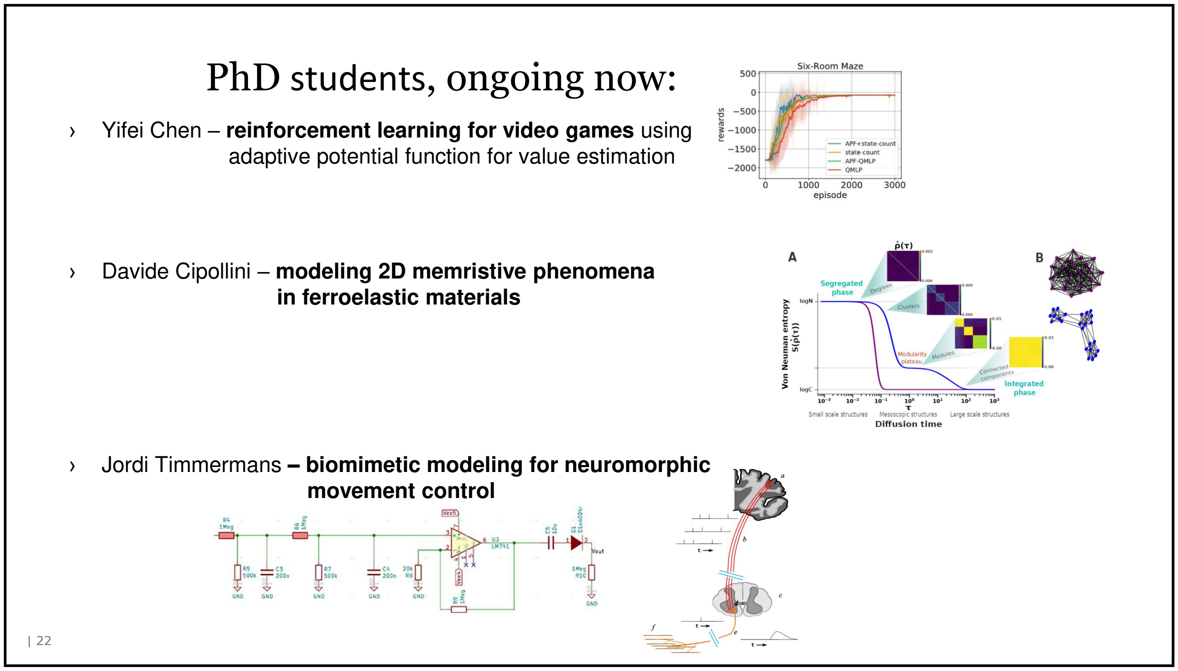
slide #21

I alse have PhD students where I was more or less a 'pro forma' or second promoter, PhD students where of course you have to look at the dissertation and see the real daily supervisors who are mentioned here in the rightmost column.



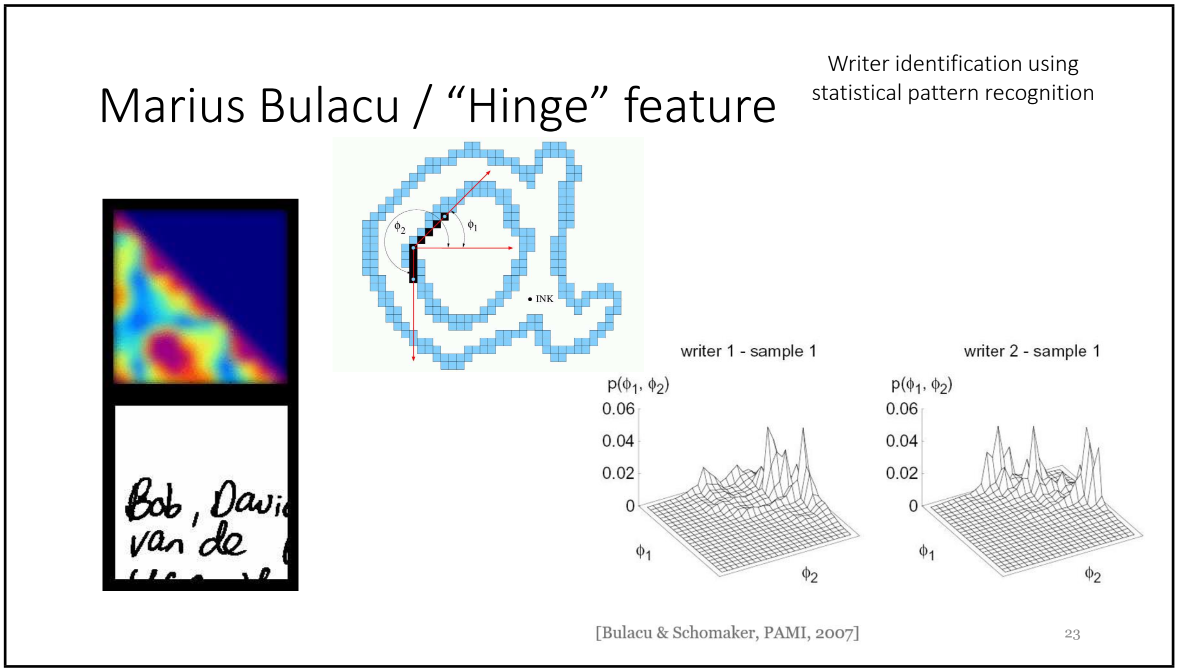
slide #22

Then, I fortunately still have three PhD students: Yifei Chen, you saw in the video; Davide Cipollini and Jordi Timmermans who are still there. With Davide I study modeling of neuromorphic materials, physics things. And, lo and behold, phase transitions are here appearing again, this is about Von Neumann multiscale entropy and graph structures. This research is in the context of the CogniGron center of neuromorphic computing and also with Jordi Timmermans you see coming back some work from my Nijmegen or Tilburg times where the brain and muscle control is translated into electronics.



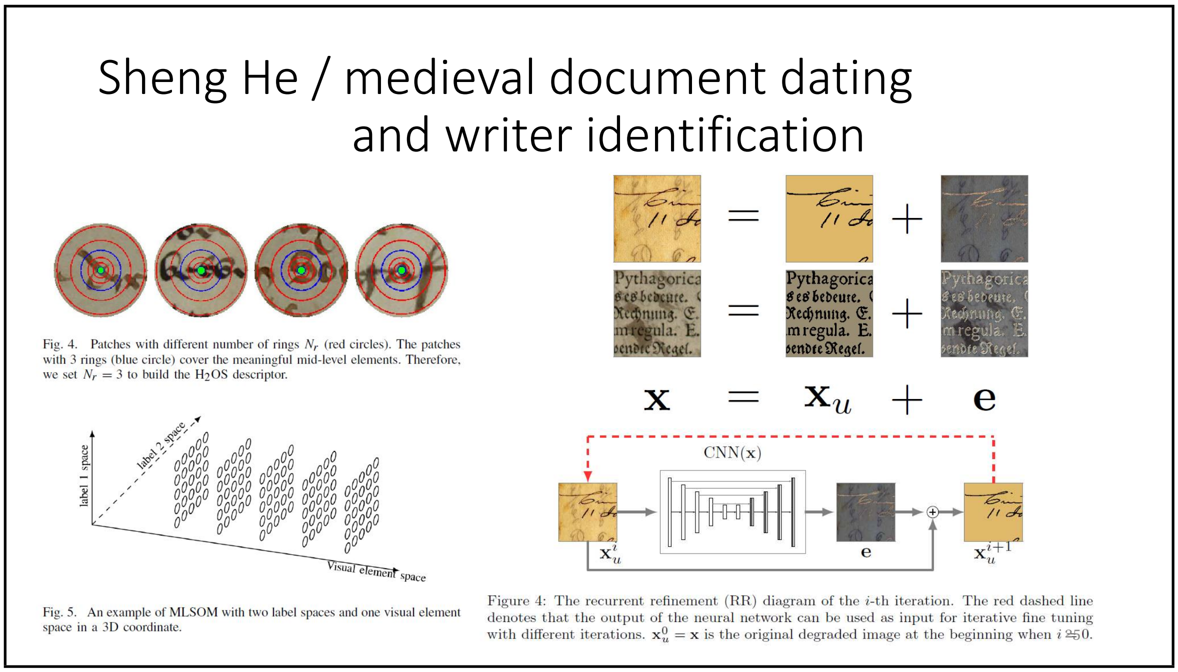
slide #23 - selected PhD studies - the 'Hinge' (Marius Bulacu)

Marius Bulacu already had his presentation this morning. There is a nice movie (animated gif) on this slide that he didn't show but you can see that if you take this kind of landscape of probabilities of angle combinations along handwriting that for each of those writers, there are several writers, you can see that their landscape of densities is different for each writer and this is why we also, in the end, were able to use it in the Dead Sea Scrolls project with Mladen Popovic.



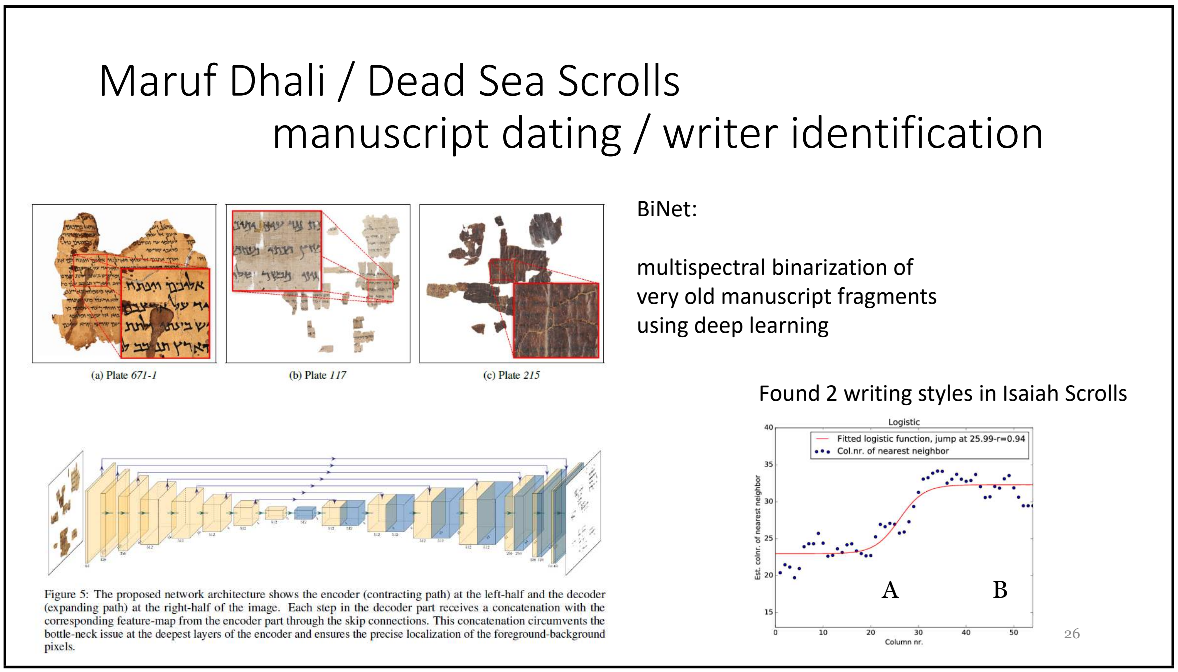
slide #25 - DeepOtsu and FragNET (Sheng He)

Sheng He did also many interesting things. In this case, for instance, he made a neural network architecture to separate the complex textured paper background from ink in a slightly different way than Maruf Dhali did in BiNET. Sheng also developed a model for time-axis modeling where you can have different style fragments of handwriting along a time axis.



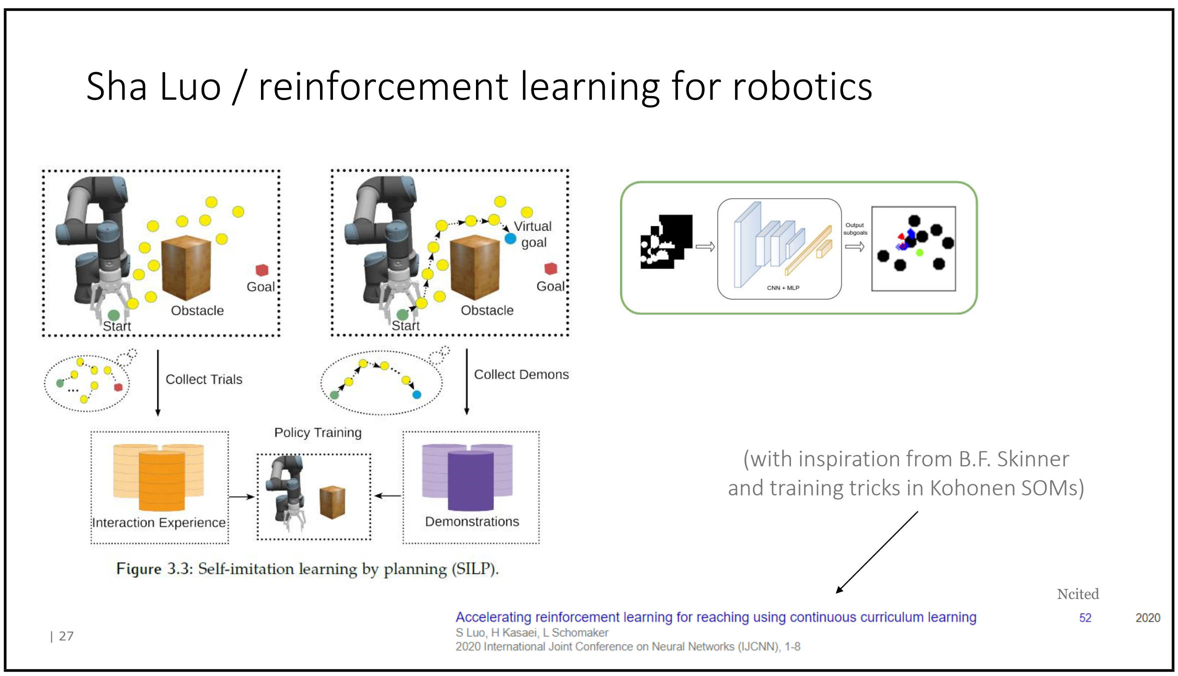
slide #26 - Discovery of two writers in Isaiah Scrolls (Maruf Dhali)

Maruf already told about this in his presentation. Here you can see the phase transition, on this slide, in our PlosOne article on the Isaiah scrolls, with a transition in writing style from writer A to writer B. This result is also one of the reasons why I used 'phase transitions' as the general theme for this presentation.



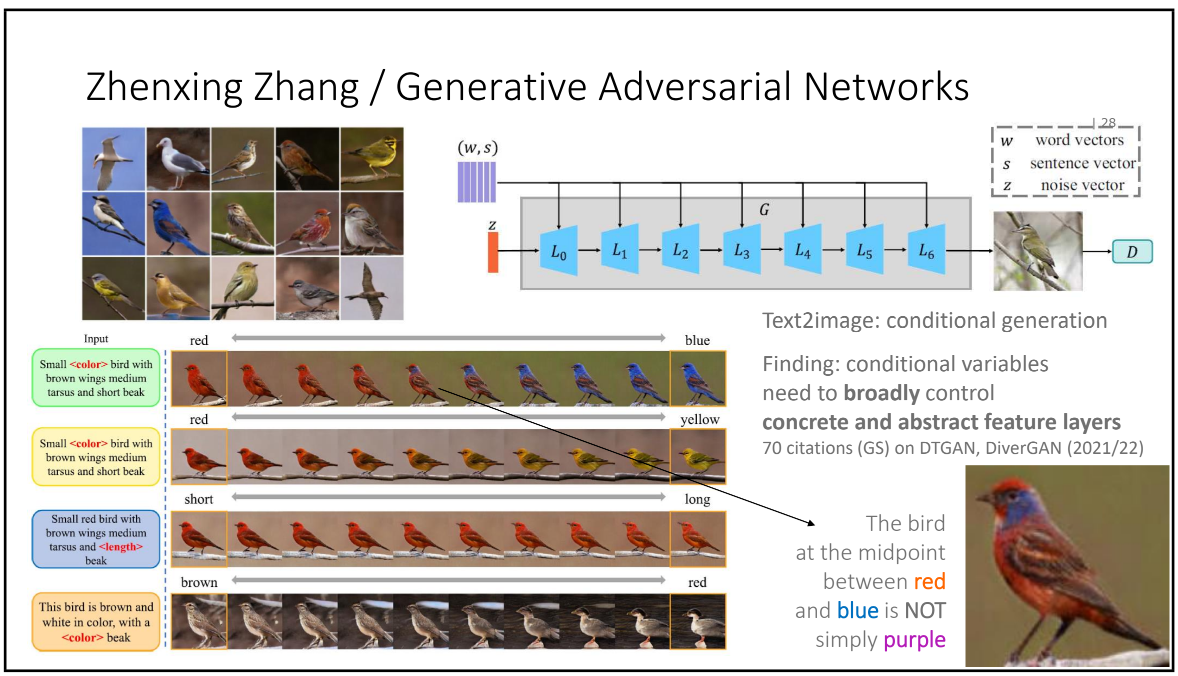
slide #27 - Robotics and reinforcement learning (Sha Luo)

Sha Luo was also there, in the video, this morning. Robotics we also like a lot but reinforcement learning is very expensive in deep learning so for for a video game, for instance, you need 10 million video frames before the system learns it. In robotics you need to do this in simulated space and here you can see that obstacle avoidance can be trained with reinforcement learning. We do this in a kind of mixture between traditional movement planning, i.e., algorithms in the robot that implement the inverse kinematics task, and we train the neural network with that. Actually this is a mixture between reinforcement learning and supervised learning. Another thing that is cited quite a lot (52 times already) is a trick that's from my psychology background so I know how BF Skinner trained his pigeons and I also reused some training tricks from my variant of the Kohonen self-organized maps. That means that the teacher in the beginning is lenient and then we have a curve which makes the requirements gradually more strict and of course the question is whether you do this with a convex or a concave curve, namely, is the teacher becoming strict very quickly or is the teacher lenient and doing it slowly. This is an idea that apparently is picked up in in literature quite well.

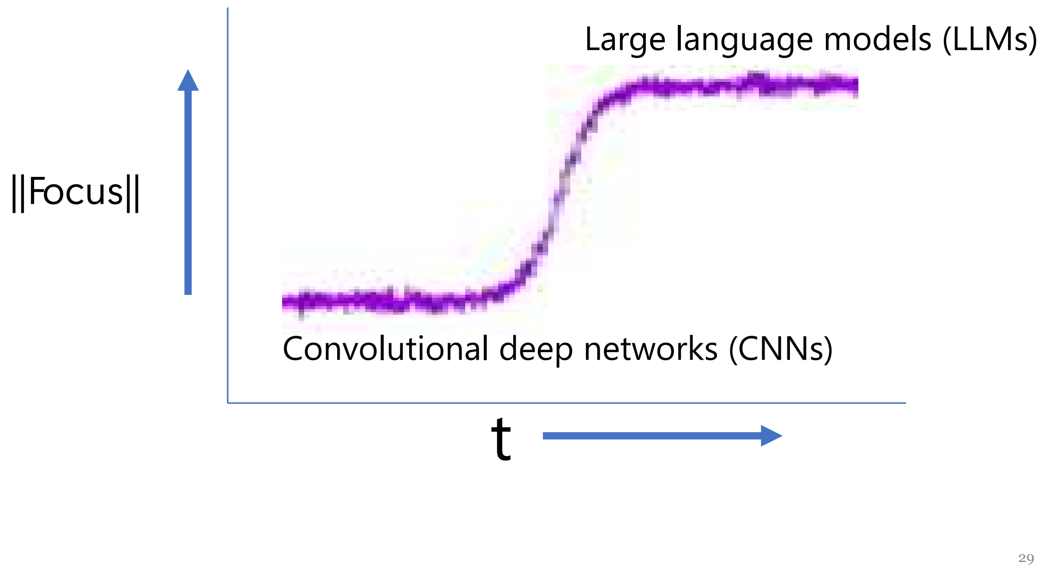


slide #28 - Generative Adversarial Networks (Zhenxing Zhang)

This is another very nice example, this is not about deep-faking Trump but this is about typing sentences and then having photographs out. The data set here concerns birds and then the system, the generative adversarial network (GAN) needs to produce a believable bird image in the end. It's too complicated to tell all the details but two things I can say. The first thing is that if you want to influence the rendering of these complex images the text information that describes the type of bird, for instance "small red bird with brown wings medium size and short beak" that kind of conditioning information cannot be injected in one place in such a neural network, you have to spread it out over the concrete and abstract shape layers in the the representation, so that the total image in the end is influenced by the words in that text. I have to speed up but if you would have said to me before this research: "What would happen if you have a latent space and there is a kind of believable but non-existing red bird and I have some kind of non-existing blue bird generated by the GAN, if you would have asked me what is the bird in the middle then with the usual computer-vision background you would say hmm, red and blue: Even Elin (grandchild of 6yrs, present), knows that if you mix that in RGB space, the middle it will be purple, right? This is not what the network is doing. It is much smarter: Because of this influence at the different feature levels you can see that the transition between red and blue ends up in the middle here, with a more or less believable bird that has clear red and blue patches. So the average is not purple, the average is in a higher dimensional space z here, that describes the detailed visual features in there. The same holds for other transitions in that latent space.



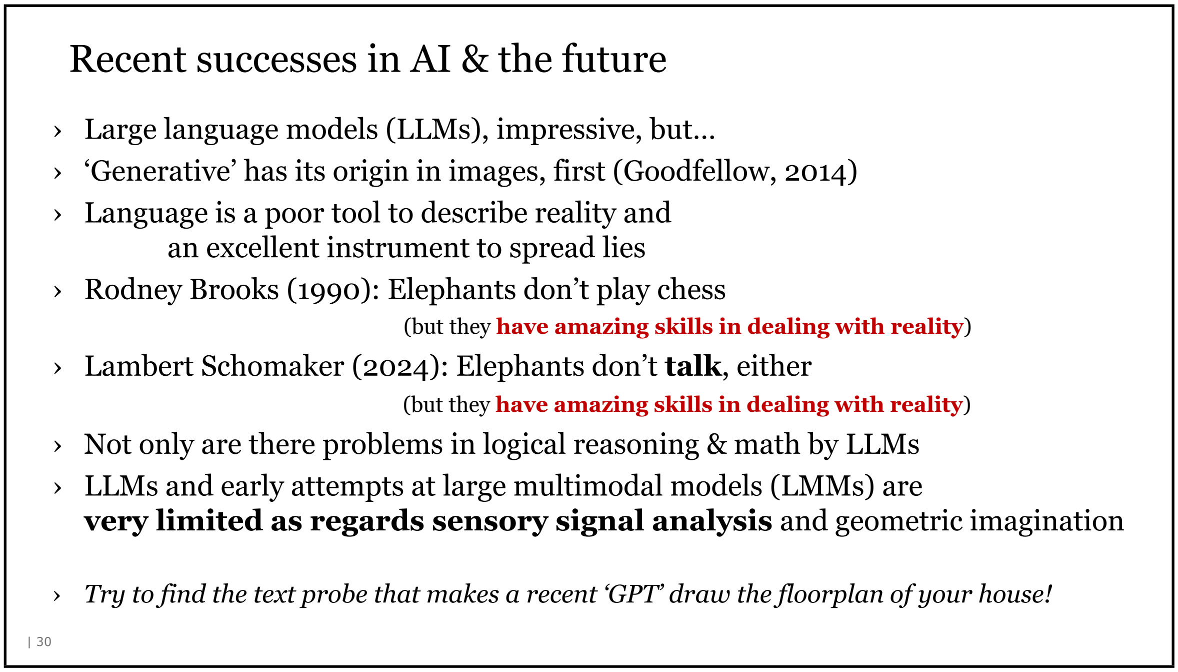
slide #29



Then the next phase transition is from the convolutional neural networks (CNNs, the ones by Yann LeCun who was mentioned before) to the large language models (LLMs) and here I can say a little bit more and so of course these language models are very impressive. You have seen examples of it, of course you have also seen the terrible mishaps that that they do, but my irritation quickly also started when the Wikipedia site on generative AI was was introduced. Because for me, generative AI has its origin in pixel images: GANs, Goodfellow (2014) right? but that's just old school, apparently, so this whole Wikipedia website about generative AI didn't even mention Goodfellow, can you believe that? It's absolutely shocking. I think by now they have repaired it but it's quite quite terrible. I have a problem with these large language models for the following reason, and this is a very crude remark that I make but I still strongly believe in it.

> "Language is a poor tool to describe reality and an excellent instrument to spread lies"

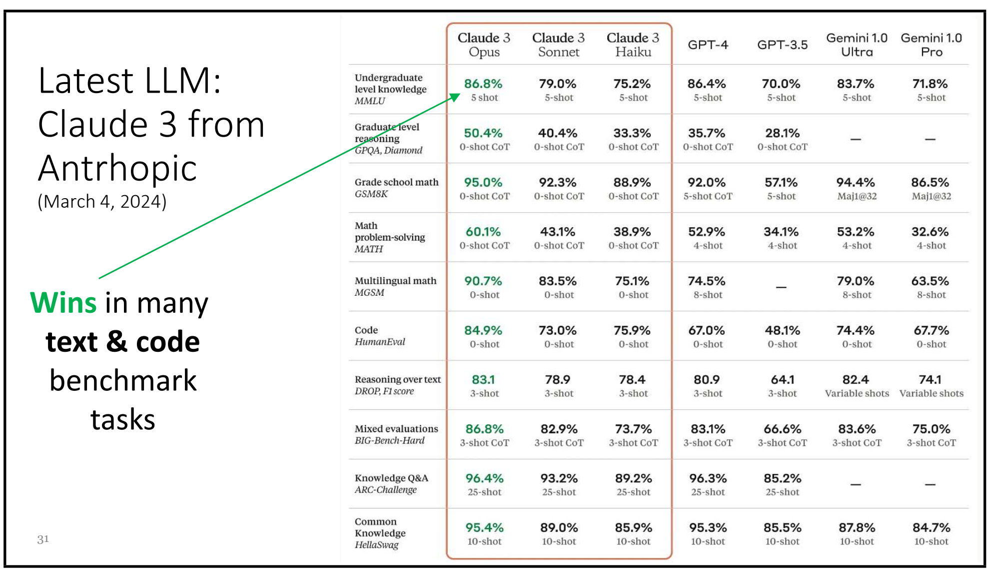
Try to tell someone over the phone how a particular sunset or sunrise was or how beautiful some waterfall looks in the Alps, or try to explain someone over the phone how to change tires on a car. Language only works because the sender and the receiver have a very big data set, so that small tokens can evoke an association. The real information is NOT in the word. So how can you expect artificial intelligence to be human like, if you only look at surface text strings? I have very many doubts about that. Already in 1990, Rodney Brooks said "Elephants don't play chess". He is from Robotics and he also disliked the idea that that intelligence is something with the game rules of chess. He had a very influential paper on that. Today I want to repeat it with the next sentence: "Elephants don't talk either"! But they have amazing skills in dealing with reality.



So, not only are there problems in logical reasoning and math by Large Language Models that Rineke Verbrugge and her group also are addressing. Now the current Large anguage Models and the early attempts at Large Multimodal Models (which are the LMMs) are very limited as regards sensory signal analysis and geometric imagination. Try to find the text prompt that makes some GPT draw the floor plan plan of your house. So you just describe: "You enter the front door, on the left is the kitchen ..." and so on and then you say, "well, give me a drawing of that floorplan". The geometric imagination of these kind of algorithms is extremely limited. So, I'm also critical of LLMs as other people in AI are, but for other reasons. But, of course - in the text space and the code space, things may work often quite well, given enough data.

slide #31 - Claude 3 by Anthropic, March 4th

This is from this week, "Claude 3" from Anthropic, the newest LLM and it really is the winner in many many tasks: For instance, "Undergraduate level knowledge" and then "math solving problems", well it's not that good (60%) but getting somewhere and better than many of the competitors. But programming code for instance, look how good code generation is. So you ask in a sentence: "Give me a sorting program in C" or something like that. You can ask more complicated code, of course, and it will be quite good. So I have only few qualms this. I think it should be good, right, because there is a massive database containing this kind of programming code data, there's a correlational model, an attentional model used to make the predictions of the token sequence so it **should** be good at this but if you look at mathematical and visual benchmark tasks, then it's still not so impressive: "Math and reasoning" only 60%, still better than many of the others, but not really fantastic and here we see even that another model is winning. But it only achieves 80% for a kind of very simple visual task - not even real computer vision tasks but simple tasks - it is very limited. So from my point of view transformers are in their infancy as regard image processing. There are models like ViT and others but they they are trained on massive data and they are still not scale or rotation invariant and, no I'm not going into the detail, I have the tendency as you know. But I had, in the beginning, in the first initiation into these language models I had all kinds of criticisms. To mention one, also other people said: "Hmm, they cannot plan so they will not be usable in robotics". After a few months, many people in robotics, also Hamidreza Kasaei, are using it and you can ask the language model to construct a plan to make something in the kitchen, to fetch something in the house and a reasonable plan comes out. I had also other qualms, other problems with it, and the one thing you should do to get these problems out, is to mention them, because of course those companies and everyone involved will say: "Hmm, is this a problem? Then we will improve it". Even up to the programming of some kind of politically correct output that now has been shown to be to yield terrible results in one of the models (the GEMINI model, Google's image generator). So if you if you press it too hard to avoid biased output you get really, really fake output.



	Claude 3 Opus	Claude 3 Sonnet	Claude 3 Haiku	GPT-4	GPT-3.5	Gemini 1.0 Ultra	Gemini 1.0 Pro
Undergraduate level knowledge *MMLU*	86.8% 5 shot	79.0% 5-shot	75.2% 5-shot	86.4% 5-shot	70.0% 5-shot	83.7% 5-shot	71.8% 5-shot
Graduate level reasoning *GPQA, Diamond*	50.4% 0-shot CoT	40.4% 0-shot CoT	33.3% 0-shot CoT	35.7% 0-shot CoT	28.1% 0-shot CoT	—	—
Grade school math *GSM8K*	95.0% 0-shot CoT	92.3% 0-shot CoT	88.9% 0-shot CoT	92.0% 5-shot CoT	57.1% 5-shot	94.4% Maj1@32	86.5% Maj1@32
Math problem-solving *MATH*	60.1% 0-shot CoT	43.1% 0-shot CoT	38.9% 0-shot CoT	52.9% 4-shot	34.1% 4-shot	53.2% 4-shot	32.6% 4-shot
Multilingual math *MGSM*	90.7% 0-shot	83.5% 0-shot	75.1% 0-shot	74.5% 8-shot	—	79.0% 8-shot	63.5% 8-shot
Code *HumanEval*	84.9% 0-shot	73.0% 0-shot	75.9% 0-shot	67.0% 0-shot	48.1% 0-shot	74.4% 0-shot	67.7% 0-shot
Reasoning over text *DROP, F1 score*	83.1 3-shot	78.9 3-shot	78.4 3-shot	80.9 3-shot	64.1 3-shot	82.4 Variable shots	74.1 Variable shots
Mixed evaluations *BIG-Bench-Hard*	86.8% 3-shot CoT	82.9% 3-shot CoT	73.7% 3-shot CoT	83.1% 3-shot CoT	66.6% 3-shot CoT	83.6% 3-shot CoT	75.0% 3-shot CoT
Knowledge Q&A *ARC-Challenge*	96.4% 25-shot	93.2% 25-shot	89.2% 25-shot	96.3% 25-shot	85.2% 25-shot	—	—
Common Knowledge *HellaSwag*	95.4% 10-shot	89.0% 10-shot	85.9% 10-shot	95.3% 10-shot	85.5% 10-shot	87.8% 10-shot	84.7% 10-shot

Latest LLM: Claude 3 from Antrhopic

(March 4, 2024)

Mathematical and Visual benchmark tasks: *Not so impressive!!*

models. They can process a wide range of visual formats, including photos, charts, graphs and technical diagrams. We're particularly excited to provide this new modality to our enterprise customers, some of whom have up to 50% of their knowledge bases encoded in various formats such as PDFs, flowcharts, or presentation slides.

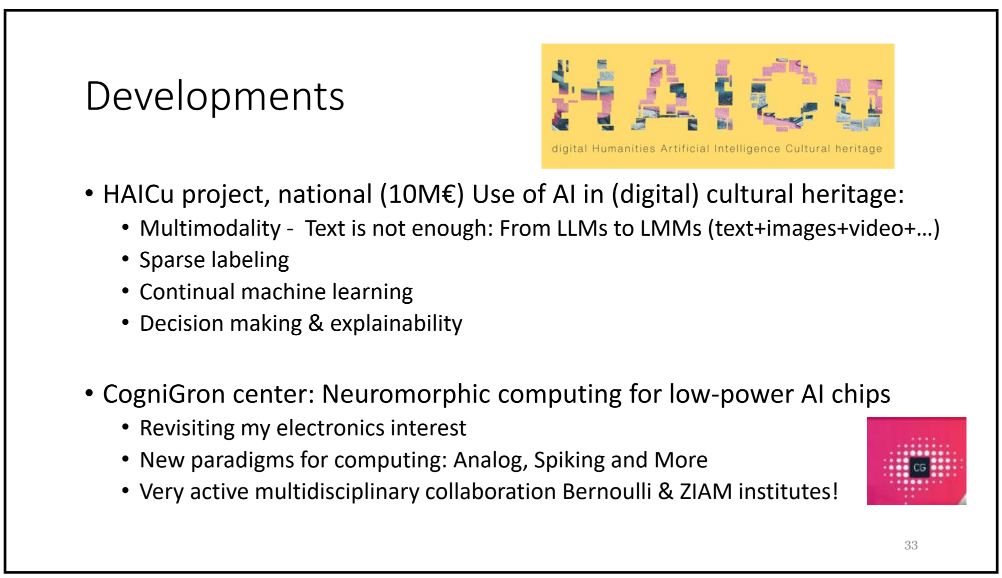
	Claude 3 Opus	Claude 3 Sonnet	Claude 3 Haiku	GPT-4V	Gemini 1.0 Ultra	Gemini 1.0 Pro
Math & reasoning *MMMU (val)*	59.4%	53.1%	50.2%	56.8%	59.4%	47.9%
Document visual Q&A *ANLS score, test*	89.3%	89.5%	88.8%	88.4%	90.9%	88.1%
Math *MathVista (testmini)*	50.5% CoT	47.9% CoT	46.4% CoT	49.9%	53.0%	45.2%
Science diagrams *AI2D, test*	88.1%	88.7%	86.7%	78.2%	79.5%	73.9%
Chart Q&A *Relaxed accuracy (test)*	80.8% 0-shot CoT	81.1% 0-shot CoT	81.7% 0-shot CoT	78.5% 4-shot CoT	80.8%	74.1%

A lot to be improved! Transformer methods are in their infancy as regards image processing but everything is moving fast.

32

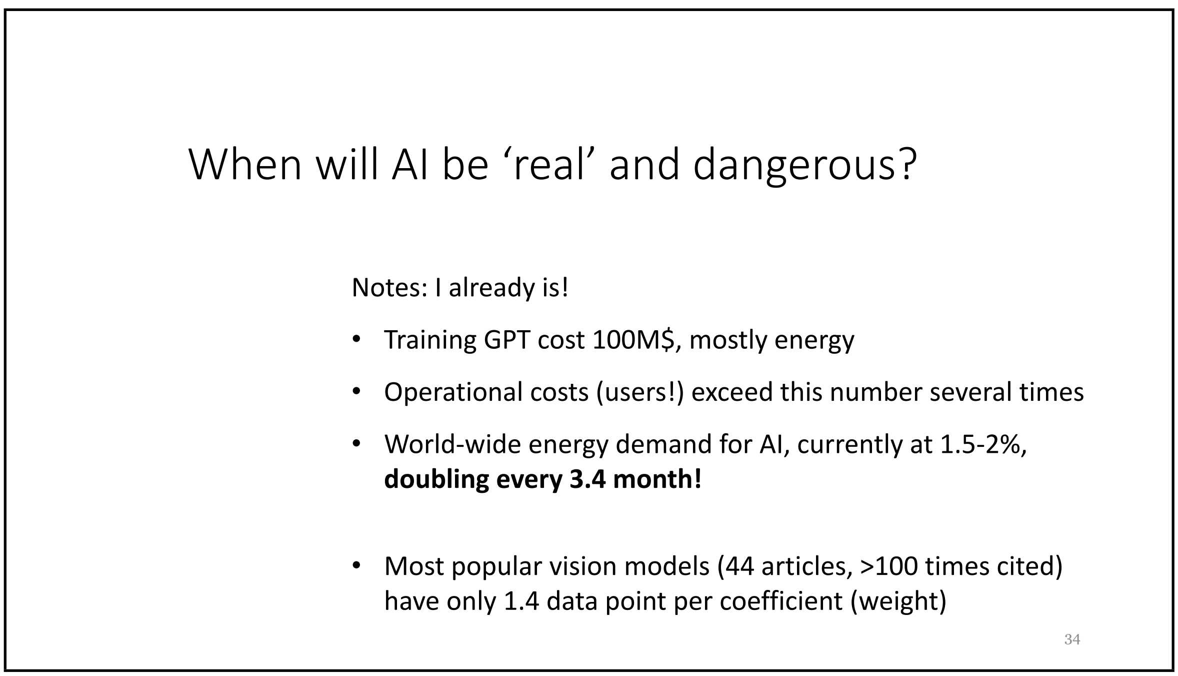
slide #33

The developments are that, indeed as was mentioned, I'm still involved in a large project applying deep learning in the cultural heritage, in these massive archival collections that exist. There are work packages in there and although the proposal was written two years ago, when the committee looked at the proposal all our themes were still relevant and are relevant today in large language models. Text is not enough. You need to go larger - look at all our senses, I I'll have a slide on that in a minute. Sparse labeling: Initially there are no examples. I mean the teacher in primary school just mentioned a few times that the letter has a particular identity and then you know it, right? Humans don't need to hear it 10,000 times and then maybe still don't know it. This is about continual machine learning. Our learning goes on all the time, we learn in real time, there is no laboratory stage of training and then an operational stage of using: Still not solved! Then: decision-making & explainability are big challenges. And then, this has to do with the next point: (picks frame with total amount for the Justdigit Foundation) has to do with this: I asked you to give a donation which has reached 717 Euro 50 for Justdigit. This is my tiny attempt to do something for the planet, because of all the bad use of energy in AI. We are making other efforts in Groningen as well, and one of it is the CogniGron center that is run by Beatriz Noheda who's also present here. What you want to have is low power AI chips that can do AI for the fraction of the energy. In a later slide I will have some numbers on that. This allows me to revisit my electronics interest, looking at analog computing, spiking and I am proud of the University of Groningen for having a center like this.



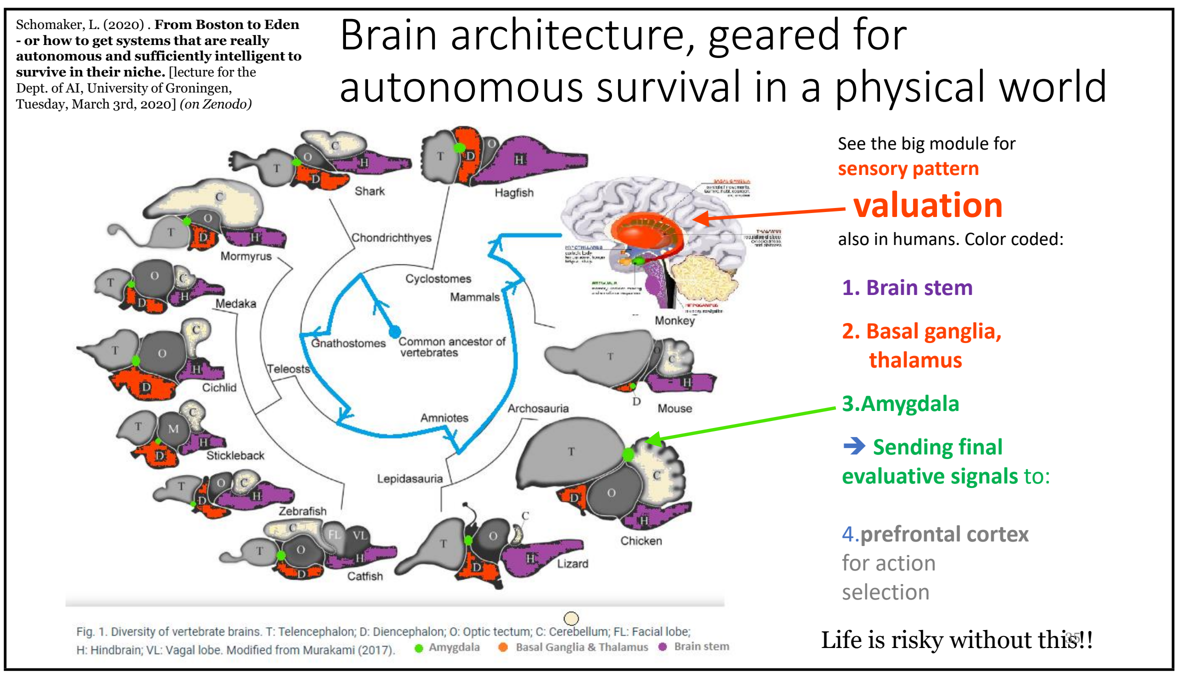
slide #34 - On current developments in AI

Now this is the question that you all be will be asking, and maybe you want to hear something about it from me: ”When will AI be real and dangerous” I would say it already is dangerous. GPT training cost $100 million, mostly energy costs but by now, because we are all using it, or many of us are using it, a multiple of this number is spent on energy alone. The worldwide energy demand for AI is currently at 1.5 to 2% and it’s doubling every 3.4 months. This is not sustainable, it is not, this is going to crash. If you talk about Nvidia stocks - I did that with my brother-in-law - then you can see if you’re smart, that there is an asymptote coming. Then another problem, that’s more scientific also very technical, but I like to go into the depth: I did a recent meta analysis of modern convolutional vision models in deep learning, on 44 articles each cited more than 100 times On average they have only 1.4 data point per weight or coefficient! That is a problem! Even if you have success, even if you do k-fold evaluation: If you don’t see that you have a problem, then there **is** a problem, right? Mathematically this is not sustainable of course. There will be subset of the data where there is some kind of say abundance of information but mathematically this is not really solved. There was only one RESNET (we talked about RESnet before, today) there was one RESNET variant that really had a solid number of data examples for training.



slide #35 - What does biological cognition have, that current AI systems do not?

So what is missing in these large language models is that, like chess, this is a kind of completely sterile world of of symbol sequences - now I go back to my physiological background (slide on brain evolution) - and if you look at the brain architecture of many species from the simple hagfish, the shark, going up to the zebra fish, the lizard, chicken and mouse and up to the human, what you can see is that there are specific brain structures that have to do with **valuation** of the input: "What is the current input for me, is it dangerous or not etc.". So you have the brain stem which is the purple box. In the brain stem the connections from all the sensors will end up it will send information to the basal ganglia that do an analysis, a very primitive analysis that old brain doesn't see all the visual details but it makes a kind of fast evaluation of the what's going on and then it all converges into this green spot here, the amydala, which then send the final evaluative signals to the prefrontal cortex so that we can do an action selection, where action also means, let's say "fight", "flight" or "take a rest". Life is risky without an architecture like this. A system that doesn't have the ability to go for food and defend itself, and detect when it's attacked etc. will not really survive. But the good news may be that current artificial intelligence also may not be so dangerous yet.



slide #36 - Dangers of AI

So let's first not forget that the most dangerous creature around will be the human, abusing computers for many years to come. There's all this fear about AI, it's a little bit strange in in a way because if you switch the energy off, then the system dies. In order to be dangerous - there's some kind of Science Fiction scenario here - the AI system needs to be both in silico, we already have that, but to be really dangerous for us it also needs to be embodied by robotics, a drone or mechatronics and it needs to be autonomous in the sense of striving to fulfill hard-coded inner needs like these biological systems have. To defend structural integrity you have to know when you're being attacked or when something is damaged and of course you need to be able to collect resources for survival like energy, CPU time, network access, you have to get it all. If I were an AI agent I would need to scour the internet for a free CPU that I could use and what you can see here is that things like

When will AI be real and dangerous?

- First: Let's not forget that the most dangerous creature around will be the human, abusing computers, for many years to come
- In order to be dangerous, AI needs to be, in my opinion:
    - Both in silico (in the computer)
      AND embodied (robotic, drone, mechatronic)
    - Autonomous in the sense of striving to fulfil hardcoded inner needs and defending structural integrity
    - Able to collect resources for survival (energy, CPU, network access)
- Several disciplines provide the tools:
  Neuroscience, **Cybernetics**, Neural networks, Robotics

36

slide #37

feedback are really coming back, like set points, concepts from cybernetics, point attractors, cyclic attractors, that are close to Henk Broer's type of work in non-linear dynamics. So you have set points for the organism and if you have an AI system that has these kind of goals and has multiple ways of achieving these, then they might become dangerous and of course my particular point is also that it should be fully multimodal: all senses, tactile, think of of all the sensors that give you pain when you hit something, right? The total skin is one big sensor for saving your structural integrity. So yeah, if you were Professor Sickbock (cartoon by Marten Toonder) not many people from other countries will know that, but there's a cartoon "Bommel" and in those stories there is a kind of Applied Professor with evil intentions, "Professor Sickbock", here's the guy, and he says: "Okay do large language model alignment post hoc, feed it with all known computer virus source codes that also contain ways of getting into your network and into your computer etc. Put that into computer systems, let it recode and improve itself and then you prompt it to survive with some criteria and you bootstrap the beast, right? Then it will be different because it is embodied, it will have some way of getting at our energy, it will have not only a way of to get into the computers, but it's a much broader risk.

When will AI be real and dangerous …

- Revival of **cybernetics principles**: Feedback, goal: set point, point attractors and cyclic attractors
- For AI: The internal state embedding should be **fully multimodal**: all senses, text, video, tactile for structural integrity both in computer space and in the physical incarnation, etc.
- *Prof. Sickbock at work:*

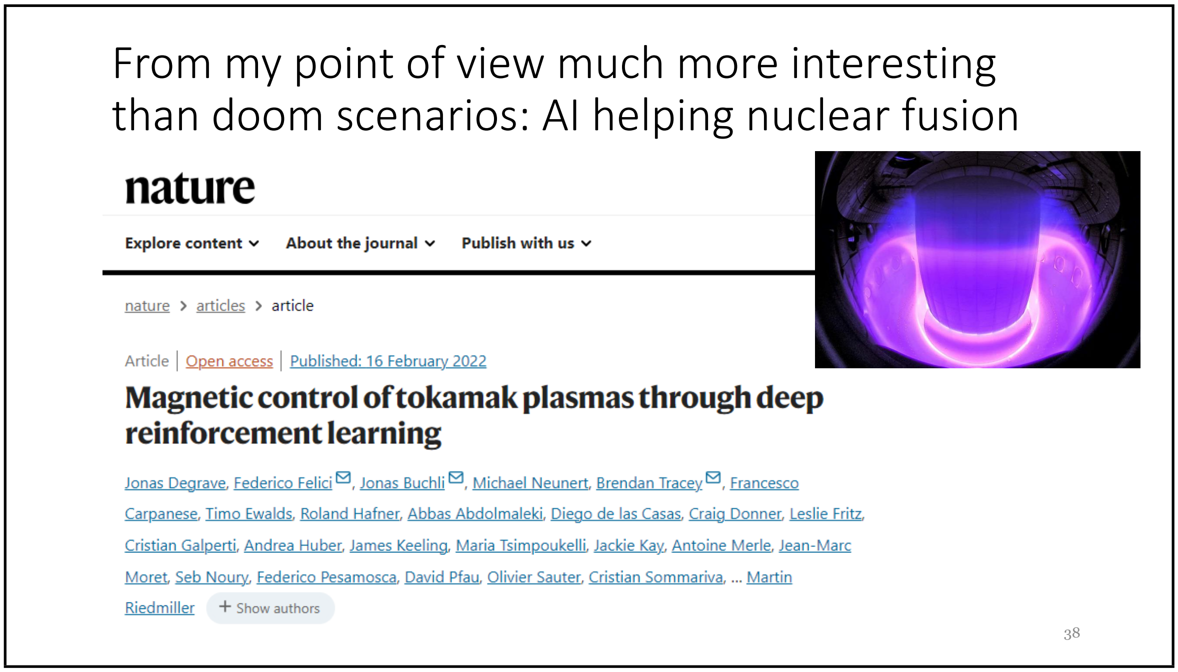
Do an LLM alignment/post tuning with all known computer-virus source codes for getting 'the AI creature' into computer systems

Prompt the survival criteria and bootstrap the beast

37

slide #38 - Reinforcement Learning useful in nuclear fusion

From my point of view, and much more interesting than doom scenarios is this, which is already two years ago, but recently it was also repeated and replicated: You can use AI to use to solve a very difficult nonlinear problem and that is to keep the plasma within a tokomak reactor from the walls to prevent damage and other big problems (cf. Egemen Kolemen/Princeton - `https://www.youtube.com/watch?v=vsc7vudaw24`) So now, the horizon for looking into the future is 300 millisecs. That looks not like a lot, but in nuclear fusion that really is a lot. - To be honest I like it better to talk about these kind of things than about dangerous AI.



Now (walks to pedestal) I have to to take a paper sheet because I don't want to make errors.

Words of thanks

At this point in my lecture, I would like to take the opportunity to say a few words of thanks. It will be impossible to mention everybody, so please don't be disappointed: I am thankful to each everybody whom I collaborated with in my career! First of all, I would like to thank my wife, Monica, who put up with the situation of living together with an eager scientist with many diverse things on his mind. It must have been difficult regularly and I thank you for keeping me grounded in the real world. Although you perfectly know the word 'toga' – for gown - you enjoyed teasing me with the phrase "so today you are wearing that kaftan again?". Then I would express my thanks and my pride in our children, Rafael and Judith, who similarly had to put up with a researcher as father. Both of you are successful in your work and private life, which may be a sign that it may not have been too much of a burden. As regards the AI department, the first person I would need to address is Elina Sietsema, who faithfully took care of all practical and organisational problems and made my life as director of the institute considerably easier. Thank you very much Elina! I already knew it beforehand, but as you gradually moved out of the picture when the AI institute was dissolved and Niels became the director of the Bernoulli Institute, I missed your continuous support. Fortunately Jan Hoogen, Sarah van Wouwe and others started to help me out. Then of course, my direct colleagues within the staff of the AI institute and later department. In spite of our culture of working together apart, there is a special cohesion among us, which was often visible and annoying to outsiders such as deans. They were never able to break us! I would especially like to thank all members of the original appointment committee who hired me and trusted me to navigate the ALICE institute through all kinds of bad weather, Niels Taatgen, Rineke Verbrugge in particular. I want to thank Fokie Cnossen for here competent steering of the educational processes. I think that by now, being a director of teaching yourself, you see that the world looks very different from above and that often our good ideas are being pulled down by the higher powers that be. I thank Bart Verheij for leading the department of AI and for the interesting discussions we had on logic versus geometry and statistics. It is important to mention Marco Wiering, who is still sadly missed. Within the APS group we had our own research interests but we had an excellent understanding and collaboration in joint PhD and MSc projects. I thank Raffaella Carloni for leading the robotics branch of my APS group. Then there are the fellow researchers and new staff of AI - like the ones I friendly called "the Mattiases", Matias Valdenegro and Matthia Sabatelli, who are taking over and will be solving a lot of the education and also doing their research – and the postdocs. I saw Gideon Maillette, also there, with several other postdocs. You also saw that more than one student got an opportunity to stay after their dissertation. You all deserve more detailed mentioning, but time is limited. Then we have the PhD students: I learned a lot from each of you! My special thanks go out to Maruf Dhali, with whom I worked on the Dead Sea Scrolls project. It was a great journey, not in the least due to the tireless scholarly and motivational input from professor Mladen Popovic, who has become a dear esteemed colleague during joint work since 2009 already. Talking about multidisciplinary collaboration, it is important to mention my colleagues in CogniGron: Ton Engbersen, from IBM with a lot of historical overlap, Tamalika Banerjee, Beatriz Noheda, the director of CogniGron, the center focusing on neuromorphic computing, where I am allowed to talk about electronics again. Many are not mentioned, I am sorry for that, so I repeat my general statement: Thank you all! I will not be completely away, having a residual appointment of 1 day in the week, but in case we do not meet: Goodbye, and all the best!

References

slide 4 Philips EE (Electronics Experimentation) kits,
`https://www.hansotten.com/electronic-kits/ee-series/ee8-ee20-a20/`

Frank, H. (1964). Kybernetische Maschinen: Prinzip und Anwendung der automatischen Nachrichtenverarbeitung, Fischer Verlag

slide 5 Braitenberg, V. (1984). Vehicles: Experiments in synthetic psychology. Cambridge, MA: MIT Press. *Note: Both in Nijmegen and Groningen, students in robotics knew Braitenberg, but not Grey Walter (1953), who definitely was first in describing sensory goal orientation and complicated behavior through simple contralateral motor control. Sic transit gloria mundi.*

Boden, M.A. (1987). Artificial Intelligence and Natural Man (2nd ed.). Basic Books.

slide 6 Walter, W. Grey (1953). The Living Brain New York: W. W. Norton & Company, Inc.

Walter, G.W. (1961) The living brain. Pelican Books.

VAX/VMS Software: VMS System Software Handbook, Digital Equipment Corporation (1985).

slide 7 Symbolics Lisp Machine,`https://en.wikipedia.org/wiki/Lisp_machine`

slide 8 EU Esprit Projects "Image & Movement Understanding" P419 and "Papyrus" P5204.

Schomaker, L., Thomassen, A. J. W. M., & Teulings, H.-L. (1987). A computational model of cursive handwriting. In R. Plamondon, C. Y. Suen, & J. G. Deschenes (Eds.), Proceedings of the Third International Symposium on Handwriting and Computer Applications (pp. 5-7).

slide 10 Rumelhart, D. E., McClelland, J. L., & The PDP Research Group. (1986). Parallel distributed processing: Explorations in the microstructure of cognition: Foundations. MIT Press, Vol 1 (blue) + 2 (brown) +Student book (yellow) and diskette.

Marvin Minsky and Seymour Papert (1969). Perceptrons, Cambridge, MA: MIT Press *Talking about scientific urban legends. Of course the failed XOR mapping is an important finding. But is that sufficient to fully eradicate neural networks, as a principle? Many systems are linear. In spite of the new urban legend, the Widrow-Hoff delta rule was still happily used for echo cancellation in telephone landlines, for a long time. Lemming behavior in science!*

Schomaker, L.R.B. (1991). Simulation and recognition of handwriting movements: a vertical approach to modeling human motor behavior, dissertation, Radboud University Nijmegen (NICI). `https://repository.ubn.ru.nl/handle/2066/113914`

Schomaker, L.R.B. (1992). A neural oscillator-network model of temporal pattern generation, Human Movement Science, Volume 11, Issues 1–2, February 1992, Pages 181-192, `https://doi.org/10.1016/0167-9457(92)90059-K`

slide 11 Elman, J. L. (1990). Finding structure in time. Cognitive Science, 14(2), 179–211, `https://doi.org/10.1016/0364-0213(90)90002-E`

slide 12 Lecun, Y., Boser, B., Denker, J. S., Henderson, D., Howard, R. E., Hubbard, W., Jackel, L. D., & Baird, H. S. (1990). Constrained neural network for unconstrained handwritten digit recognition. In C. Suen (Ed.), Frontiers in Handwriting Recognition, Montreal, 1990 CENPARMI, Concordia University.

slide 13 Boser, B., Guyon, I., Vapnik, V. (1992). A training algorithm for optimal margin classifiers. In: Proceedings of the Fifth Annual Workshop on Computational Learning Theory, Pittsburgh.

slide 15 Schomaker, L. (2016). Design considerations for a large-scale image-based text search engine in historical manuscript collections. In it - Information Technology (Vol. 58, Issue 2, pp. 80–88). Walter de Gruyter GmbH. https://doi.org/10.1515/itit-2015-0049 *Description of the Monk trainable search engine, using fast trainable centroid-based word image matching with a handcrafted feature, 'quadpatch', from 2009*

slide 16 Alex Graves, Santiago Fernández, Marcus Liwicki, Horst Bunke, and Jürgen Schmidhuber (2007). Unconstrained online handwriting recognition with recurrent neural networks. In Proceedings of the 20th Int. Conf. on Neural Information Processing Systems (NIPS'07). Curran Associates Inc., Red Hook, NY, USA, 577–584.

slide 17 van Oosten, J.-P. (2021). The snowball principle for handwritten word-image retrieval: The importance of labelled data and humans in the loop. **dissertation**. University of Groningen. `https://doi.org/10.33612/diss.160750597`

slide 18 Graves, A., Jaitly, N., & Mohamed, A. (2013). Hybrid speech recognition with Deep Bidirectional LSTM. In 2013 IEEE Workshop on Automatic Speech Recognition and Understanding. 2013 IEEE Workshop on Automatic Speech Recognition & Understanding (ASRU). IEEE. `https://doi.org/10.1109/asru.2013.6707742`

slide 19 Velich, Maximilian (2022) Context is All You Need: Comparing Transformer-based to RNN-based Models in a Handwritten Text Recognition Task. Master's Thesis, Artificial Intelligence, `https://fse.studenttheses.ub.rug.nl/27947/`

slide 22 Y. Chen, H. Kasaei, L. Schomaker and M. Wiering, "Reinforcement Learning with Potential Functions Trained to Discriminate Good and Bad States," 2021 International Joint Conference on Neural Networks (IJCNN), Shenzhen, China, 2021, pp. 1-7, doi: 10.1109/IJCNN52387.2021.9533682.

Cipollini, D. & Schomaker, L.R.B. (2023). Neuromorphic Computing and Engineering, 3, 034001, DOI 10.1088/2634-4386/acd6b3 Schomaker, L., Timmermans, J., & Banerjee, T. (2023). Non-linear

adaptive control inspired by neuromuscular systems. In Bioinspiration & Biomimetics (Vol. 18, Issue 4, p. 046015). IOP Publishing. `https://doi.org/10.1088/1748-3190/acd896`

slide 23 Bulacu, M. & Schomaker, L. (2007). Text-Independent Writer Identification and Verification Using Textural and Allographic Features, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 29, no. 4, pp. 701-717

slide 24 Dijkstra, K., van de Loosdrecht, J., Atsma, W. A., Schomaker, L. R. B., & Wiering, M. A. (2021). CentroidNetV2: A hybrid deep neural network for small-object segmentation and counting. In Neurocomputing (Vol. 423, pp. 490–505). Elsevier BV. `https://doi.org/10.1016/j.neucom.2020.10.075`

slide 25 He, S., Samara, P., Burgers, J., & Schomaker, L. (2016). A Multiple-Label Guided Clustering Algorithm for Historical Document Dating and Localization. In IEEE Transactions on Image Processing (Vol. 25, Issue 11, pp. 5252–5265). Institute of Electrical and Electronics Engineers (IEEE). `https://doi.org/10.1109/tip.2016.2602078`

He, S., & Schomaker, L. (2019). DeepOtsu: Document enhancement and binarization using iterative deep learning. In Pattern Recognition (Vol. 91, pp. 379–390). Elsevier BV. `https://doi.org/10.1016/j.patcog.2019.01.025`

slide 26 Popović, M., Dhali, M. A., & Schomaker, L. (2021). Artificial intelligence based writer identification generates new evidence for the unknown scribes of the Dead Sea Scrolls exemplified by the Great Isaiah Scroll (1QIsaa). In D. R. Amancio (Ed.), PLOS ONE (Vol. 16, Issue 4, p. e0249769). Public Library of Science (PLoS). `https://doi.org/10.1371/journal.pone.0249769`

slide 27 Luo, S., Kasaei, H., & Schomaker, L. (2020). Accelerating Reinforcement Learning for Reaching Using Continuous Curriculum Learning. In 2020 International Joint Conference on Neural Networks (IJCNN). 2020 International Joint Conference on Neural Networks (IJCNN). IEEE. `https://doi.org/10.1109/ijcnn48605.2020.9207427`, `https://www.youtube.com/watch?v=WY-1EbYBSGo`

slide 28 Zhang, Z., & Schomaker, L. (2021). DTGAN: Dual Attention Generative Adversarial Networks for Text-to-Image Generation. In 2021 International Joint Conference on Neural Networks (IJCNN). 2021 International Joint Conference on Neural Networks (IJCNN). IEEE. `https://doi.org/10.1109/ijcnn52387.2021.9533527`

Zhang, Z., & Schomaker, L. (2022). DiverGAN: An Efficient and Effective Single-Stage Framework for Diverse Text-to-Image Generation. Neurocomputing (Vol. 473, pp. 182–198). Elsevier, DOI: 10.1016/j.neucom.2021.12.005

slide 30 Brooks, R. A. (1990). Elephants don't play chess. In Robotics and Autonomous Systems (Vol. 6, Issues 1–2, pp. 3–15). Elsevier BV. `https://doi.org/10.1016/s0921-8890(05)80025-9`

slide 31 Anthropic: Introducing the next generation of Claude (4th March 2024), `https://www.anthropic.com/news/claude-3-family`

slide 33 - HAICu project: digital Humanities - Artificial Intelligence - Cultural heritage,
`https://www.nwo.nl/en/projects/nwa151822105`

- CogniGron center on neuromorphic computing,
`https://www.rug.nl/research/fse/cognitive-systems-and-materials/about/`

slide 35 Schomaker, L. (2020). From Boston to Eden - or how to get systems that are really autonomous and sufficiently intelligent to survive in their niche **lecture**, `https://zenodo.org/records/3757369`

slide 38 Degrave, J., Felici, F., Buchli, J., Neunert, M., Tracey, B., Carpanese, F., Ewalds, T., Hafner, R., Abdolmaleki, A., de las Casas, D., Donner, C., Fritz, L., Galperti, C., Huber, A., Keeling, J., Tsimpoukelli, M., Kay, J., Merle, A., Moret, J.-M., . . . Riedmiller, M. (2022). Magnetic control of tokamak plasmas through deep reinforcement learning. Nature (Vol. 602, Issue 7897, pp. 414–419). Springer Science and Business Media LLC. `https://doi.org/10.1038/s41586-021-04301-9`

The video and Powerpoint slides of this presentation can be found on
`https://zenodo.org/records/10967461`.