No SUMMA līdz dBaby projektam

Guntis Bārzdiņš, Didzis Goško, Renārs Liepiņš

LU 76. konference

Datorlingvistikas sekcija

01/03/2018

IMCS, University of Latvia (AI Lab)

H2020 SUMMA projekts

Guntis Bārzdiņš
guntis.barzdins@leta.lv
LETA Innovation Labs

http://summa-project.eu
ICT-16-2015 - Big data – research (01/02/2016 – 31/01/2019)

Total Costs in the Proposal: 9,858,326.25 € **(Latvijai 1,16 MEUR)**

BigData: Scalable Understanding of Multilingual MediA (SUMMA)

Media Item | Trending Stori

https://ui.summa.leta.lv/#/trending/all/stories/22/media-items/694298ac-f2cc-4a2b-911a-9ba78dde170b

Trending

Entities

Logout

Logged in as guntis3

Feedback

Media Item

Create a new query +

Queries > Story: Australia - 26 years without a recess… > **Media Item: Australia - 26 years without …**

Australia - 26 years without a recession.

Added: 2 hours ago (2017-06-30 08:50 UTC)
Changed: 2 hours ago (2017-06-30 08:53 UTC)
Source: DW German Videos

Transcript:

Original Live (de) | English | Original (de)

ob es sie schon bei dem es zwischen m das herumgesprochen hat aus doch in der löst im rekord der in die sauer im jahr neunzehn hundert ein und neunzig gab es den letzten ausreißer nach unten die wirtschaft doch seit man sechsundzwanzig yaam deckt die wirtschaft zu um einen mi amal weniger stark aber immer im plus australien ist mit bodenschätzen reich gesegnet große mengen von eisenerz gas öl und kohle werden von australien nach china in den immobilie er inzwischen wo man dann sind teuer die regierung tritt bei der einwanderungspolitik deshalb schon auf so wird jetzt flüssiges englisch vorausgesetzt das ist dass der schlüssel für eine erfolgreiche integration in die australische gesellschaft ist für wirtschaftlichen und sozialen ab wolk um teil der australischen gemeinschafts werden muss man englisch beherrschen die büro in australien wachsen kaum noch aus den spielraum für konsum sehr stark einschränkt im vertrauen auf grenzenloses wachstum haben sich viele

Trending
Entities
Logout

Media Item

Create a new query +

Queries > Story: Australia - 26 years without a recess… > **Media Item: Australia - 26 years without …**

Australia - 26 years without a recession.

Added: 2 hours ago (2017-06-30 08:50 UTC)
Changed: 2 hours ago (2017-06-30 08:53 UTC)
Source: DW German Videos

Transcript:

Original Live (de) | English | Original (de)

Whether it is already spoken of in between m, but in the record in the record of the sauer one hundred and ninety hundred and ninety-nine hundred and ninety-four there were the last Aussie to the bottom of the economy .but since becoming six-and-20 Yaam covers the economy to a mi Atimes less strong, but always in the plus .**Australia** is richly blessed with land treasures, large amounts of iron ore, gas, oil and coal are made by australs to china .into immobilia he by now .where you then are expensive .so the government is already appearing on immigration policy .that's how liquid English is now provided .this is that the key to successful integration into the Australian society is for economic and social ab Wolk around the Australian community, you must master English .the office in Australien is hardly growing, the scope for consumption very much in the confidence on borderline growth many australia high-debt have found themselves in the midst of the joy about the growth record had to be critical and they showed unrestrained .could be in austere Balz.

SUMMA Platform Advantages

- Completely self-contained
  - No dependence on external (cloud) services
  - All components developed within SUMMA by top Kaldi, WMT, SemEval contributors
  - 9 languages: EN, DE, AR, RU, SP, LV, PT, UR, IR
- Scalable for BigData
  - All NLP modules are Docker containers
  - Scalable to 400 live streams on 800 servers (e.g. AWS)
  - No external licencing

examples of H2020 LT projects

Call 1 ICT-17 projects on machine translation

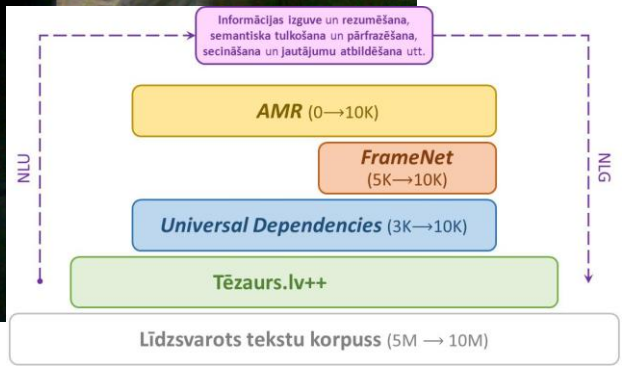
- MMT: adaptive domain-sensitive MT infrastructure
- QT21: SoA neural MT technology
- TraMOOC: Translation for Massive Open Online Courses
- HimL: automated translation of public health information

Examples of **H2020 Big Data projects with LT dimension**

- FREME: semantic enrichment of digital content
- Kconnect: ML search for medical information
- SUMMA: media analytics Big Data platform - automated analysis of media streams across many languages

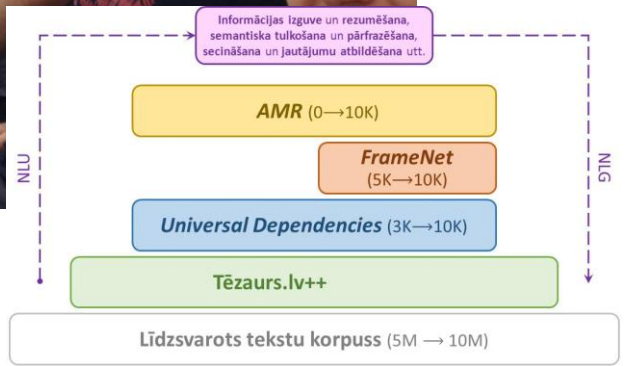
Mazais Ansis 2008 → AILab 2017 → …

SEMANTICS	deep	**MODEL**	FOL (ontology), modality, scope, planning, reasoning, "discourse"
		COREFERENCE	Anaphoras, named entities Event structure
	shallow	**FRAMENET**	"Who did what to whom" Coarse-grained word senses
GRAMMAR		**SYNTAX**	Dependency links Phrase structure
		MORPHOLOGY	Lemmas POS tags



Mazais Ansis 2017: KociņšBurgerī un Tekstrade

SEMANTICS	deep	**MODEL**	FOL (ontology), modality, scope, planning, reasoning, "discourse"
	shallow	**COREFERENCE**	Anaphoras, named entities Event structure
		FRAMENET	"Who did what to whom" Coarse-grained word senses
GRAMMAR		**SYNTAX**	Dependency links Phrase structure
		MORPHOLOGY	Lemmas POS tags



Presentation of Proposal idea

digital Baby (dBaby)

Towards super-human NLP quality : Visually Grounded Language Learning

Guntis Barzdins
guntis.barzdins@leta.lv
LETA Innovation Labs
leta.lv

ICT Proposers' Day
ICT-26-2018-2020: Artificial Intelligence
ICT-29-2018: A multilingual Next Generation Internet
Budapest, 10/11/2017

LETA
latvian information agency

The State-of-Art in AI

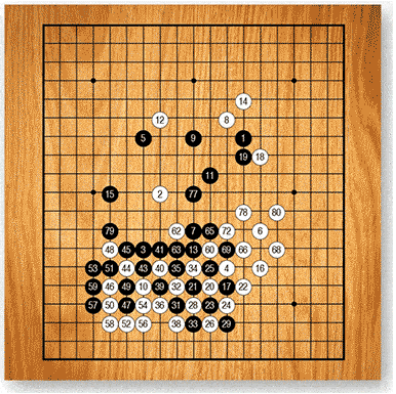
Sub-human quality
Natural Language Processing (NLP)

- Interannotator agreement for training corpora ~85%
- Supervised learning from training corpora ~70% for the complete NLP pipeline*
  - ASR (WER ~15%)
  - MT (BLEU ~30%)
  - NER / NEL (F1 ~85%)
  - AMR (SMATCH ~65%)
  - KBP (TAC KBP ~30%)

Human 100%

Super-human quality
Reinforcement Learning (RL)

- AlphaGO Zero ~110%♣



- 49 Atari games benchmark♦ (majority above human)

**Our experience in H2020 BigData-Research project SUMMA: summa-project.eu*
G.Barzdins, et. al. Character-Level Neural Translation for Multilingual Media Monitoring in the SUMMA Project. LREC 2016, pp. 1789--1793.

♣D. Silver, et al. Mastering the game of Go without human knowledge. Nature 550 (7676): 354--359 (October 2017)
♦V.Mnih, et al. Human-Level Control through Deep Reinforcement Learning. Nature 518 (7540): 529--533 (February 2015)

Achieving super-human NLP: dBaby

ATARI Reinforcement Learning framework

AlphaGO Zero RL framework (DeepMind)

dBaby: Visually Grounded Language Learning

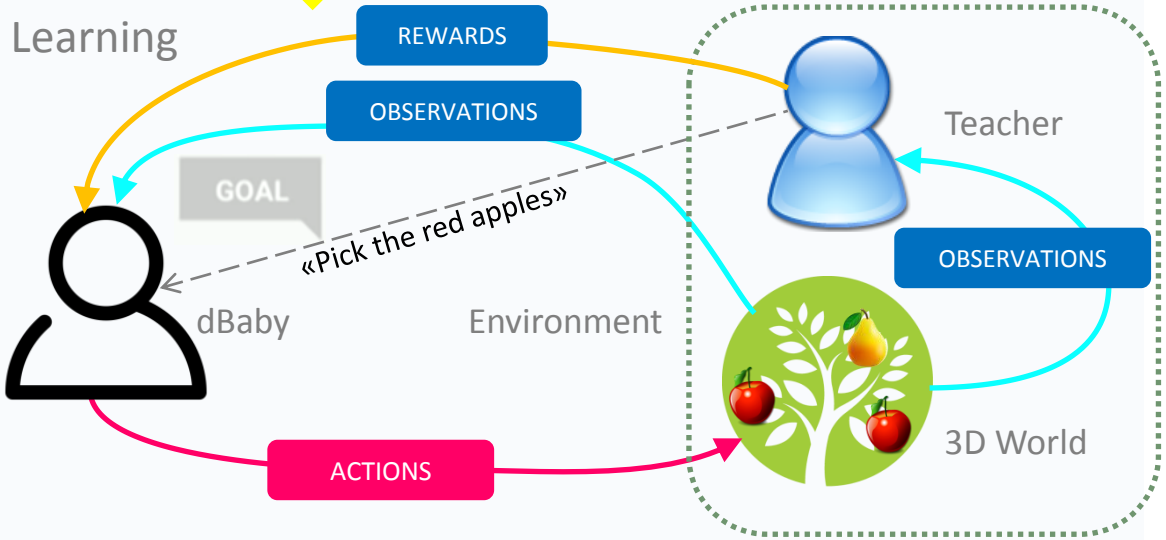
Vairs nemarķēs korpusus...

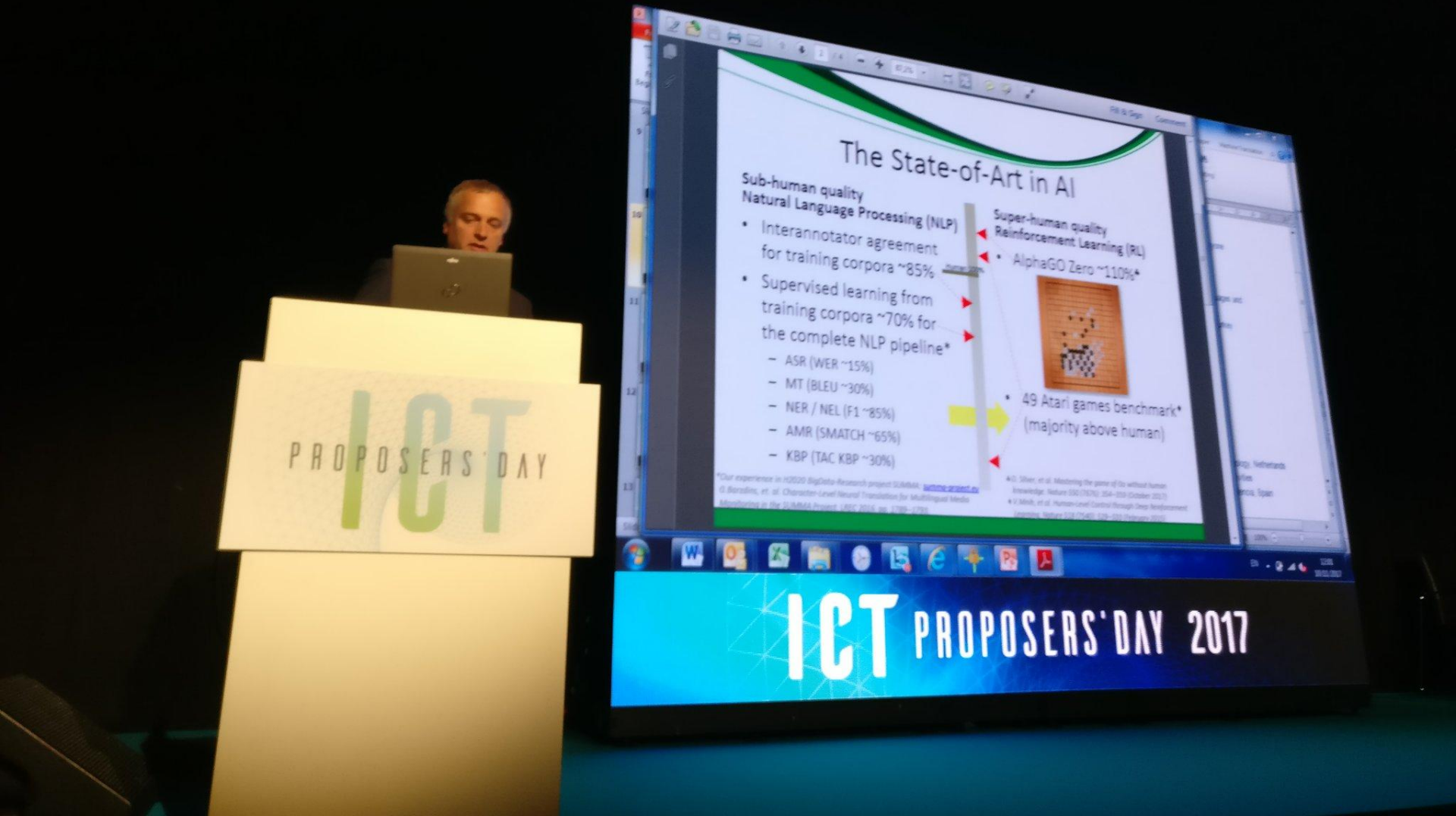
Initially a Human Teacher teaches dBaby to do what it is told to do as per:

K.M.Herman et. Al. Grounded Language Learning in a Simulated 3D World. DeepMind Research note.
https://arxiv.org/abs/1706.06551

Through auxiliary task dBaby is trained to also verbalize actions, achieving human-level NLP skills.

Afterwards two dBaby copies coordinate their actions verbally to solve tasks in ATARI RL framework, achieving super-human NLP skills.




The State-of-Art in AI
Sub-human quality
Natural Language Processing (NLP)
• Interannotator agreement for training corpora ~85%
• Supervised learning from training corpora ~70% for the complete NLP pipeline*
– ASR (WER ~15%)
– MT (BLEU ~30%)
– NER / NEL (F1 ~85%)
– AMR (SMATCH ~65%)
– KBP (TAC KBP ~30%)
Super-human quality
Reinforcement Learning (RL)
• AlphaGO Zero ~110%*
• 49 Atari games benchmark* (majority above human)
ICT PROPOSERS' DAY 2017
ICT
PROPOSERS' DAY

Neural Information Processing Systems
(NIPS), LongBeach, CA, Dec 4, 2017

How to build human-like AI?

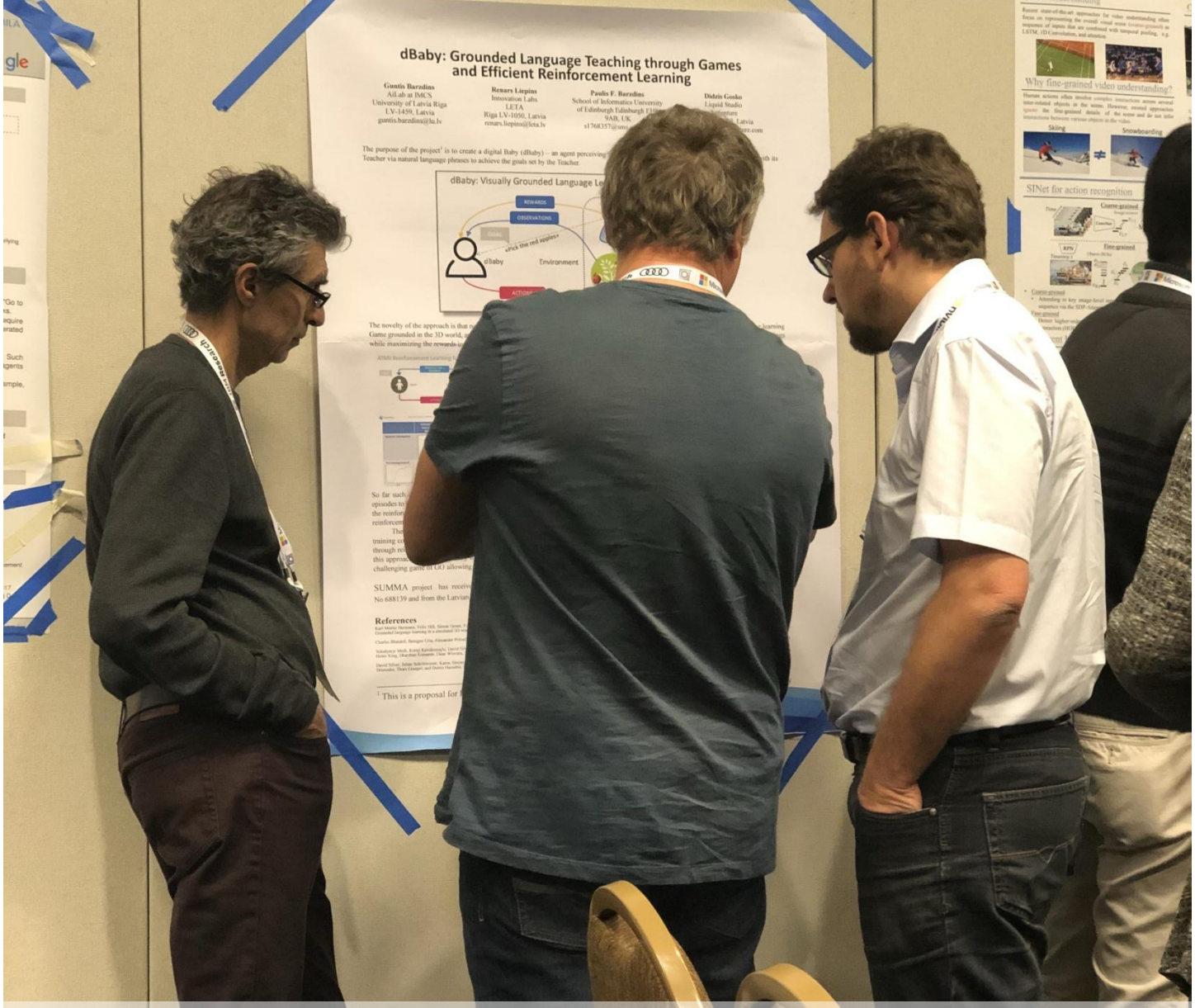
Build a machine that grows into human-level, human-like intelligence the way a person does – that starts like a baby and learns like a child.

- The oldest idea in AI, and probably the best.
- One thing Turing, Minsky, McCarthy, Hinton, LeCun, Bengio all agree on.
- It may not be the only route, but it is the one scaling route to human-level intelligence that we actually know works, reliably and (relatively) quickly.
- For the first time in history, we are in a position to take this approach seriously, so it is a good time to try it out.

Source :

How can we give robots common sense?

Josh Tenenbaum (MIT)

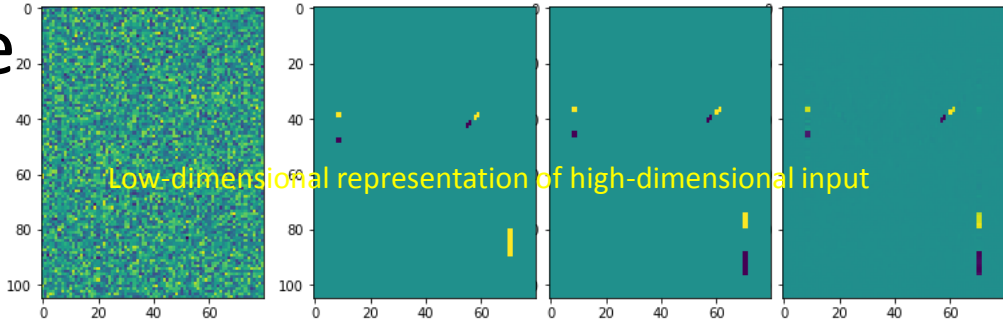


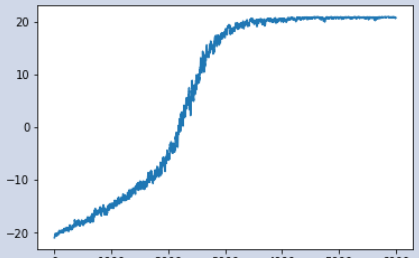
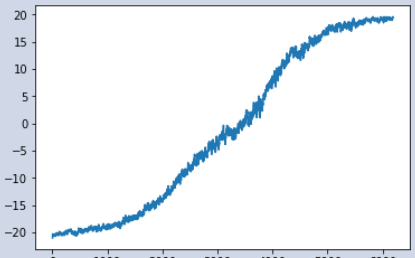
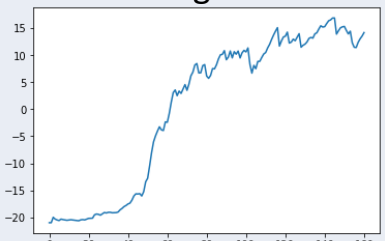
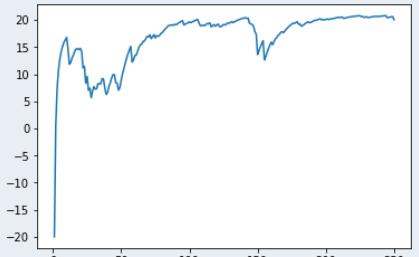
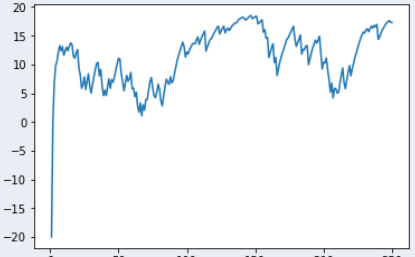
Karl Moritz Hermann, ***Felix Hill,*** *Simon Green, Fumin Wang, Ryan Faulkner, Hubert Soyer, David Szepesvari, Wojciech Marian Czarnecki, Max Jaderberg, Denis Teplyashin, Marcus Wainwright, Chris Apps,* ***Demis Hassabis,*** *Phil Blunsom. Grounded Language Learning in a Simulated 3D World.* ***DeepMind*** *Research note. https://arxiv.org/abs/1706.06551*

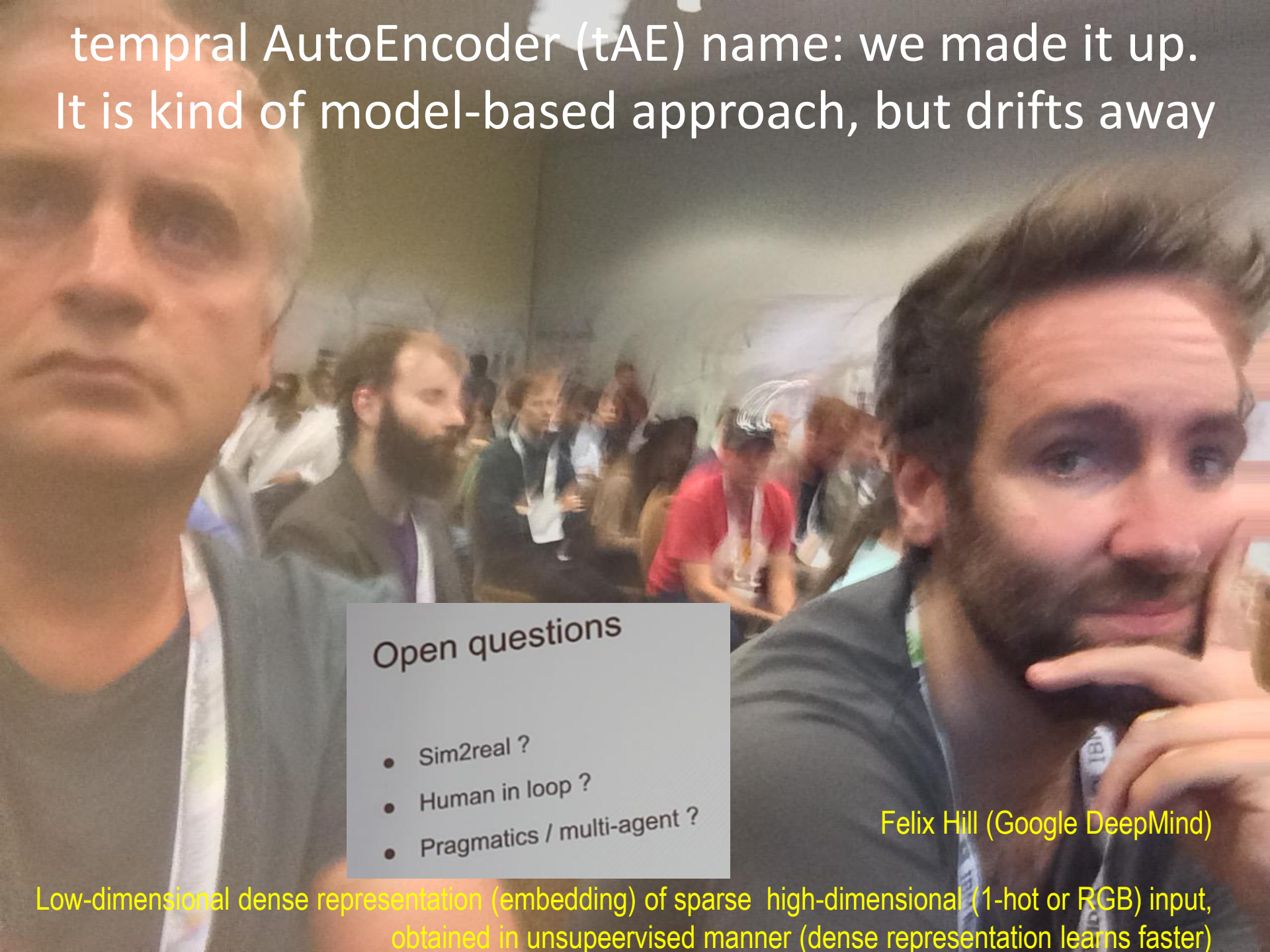
Devendra Singh Chaplot, Kanthashree Mysore Sathyendra, Rama Kumar Pasumarthi, Dheeraj Rajagopal, ***Ruslan Salakhutdinov (Apple & CMU)*** *, Gated-Attention Architectures for Task-Oriented Language Grounding, AAAI 2018. https://arxiv.org/abs/1706.07230*

G.Barzdins et.al. ***dBaby****: Grounded Language Teaching through Games and Efficient Reinforcement Learning. In: https://nips2017vigil.github.io*

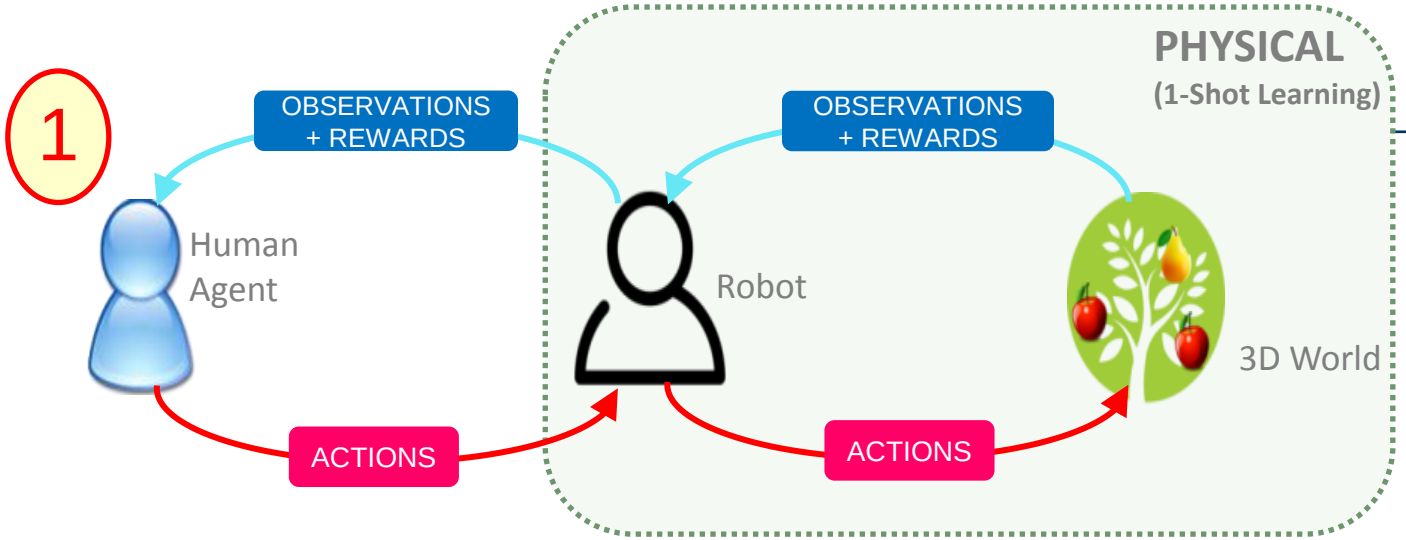
tAE: predicting the future (modeling the world)



	PolicyGradient RL from raw pixels (8400 dim)	PolicyGradient RL from tAE output (80 dim)
Random Initialization	 (a)	 (b)
Pre-training from EC	 (c)	 (d)  (e)


tempral AutoEncoder (tAE) name: we made it up.
It is kind of model-based approach, but drifts away
Open questions
Sim2real ?
Human in loop ?
Pragmatics / multi-agent ?
Felix Hill (Google DeepMind)
Low-dimensional dense representation (embedding) of sparse high-dimensional (1-hot or RGB) input,
obtained in unsupeervised manner (dense representation learns faster)

Apple Picker Roadmap

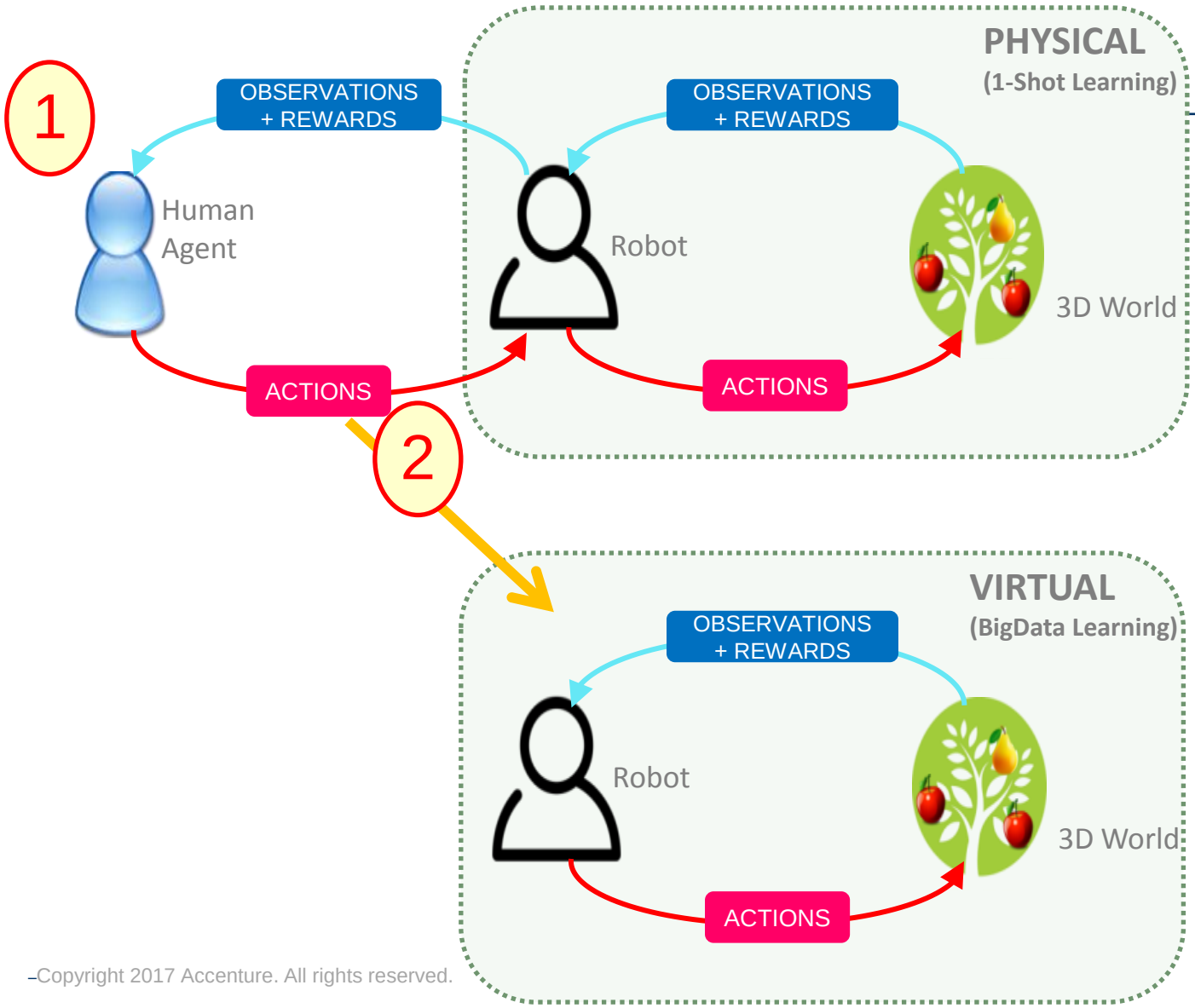


1.Human Agent controls a real Robot in real 3D world relying on video only

)

–Copyright 2017 Accenture. All rights reserved.

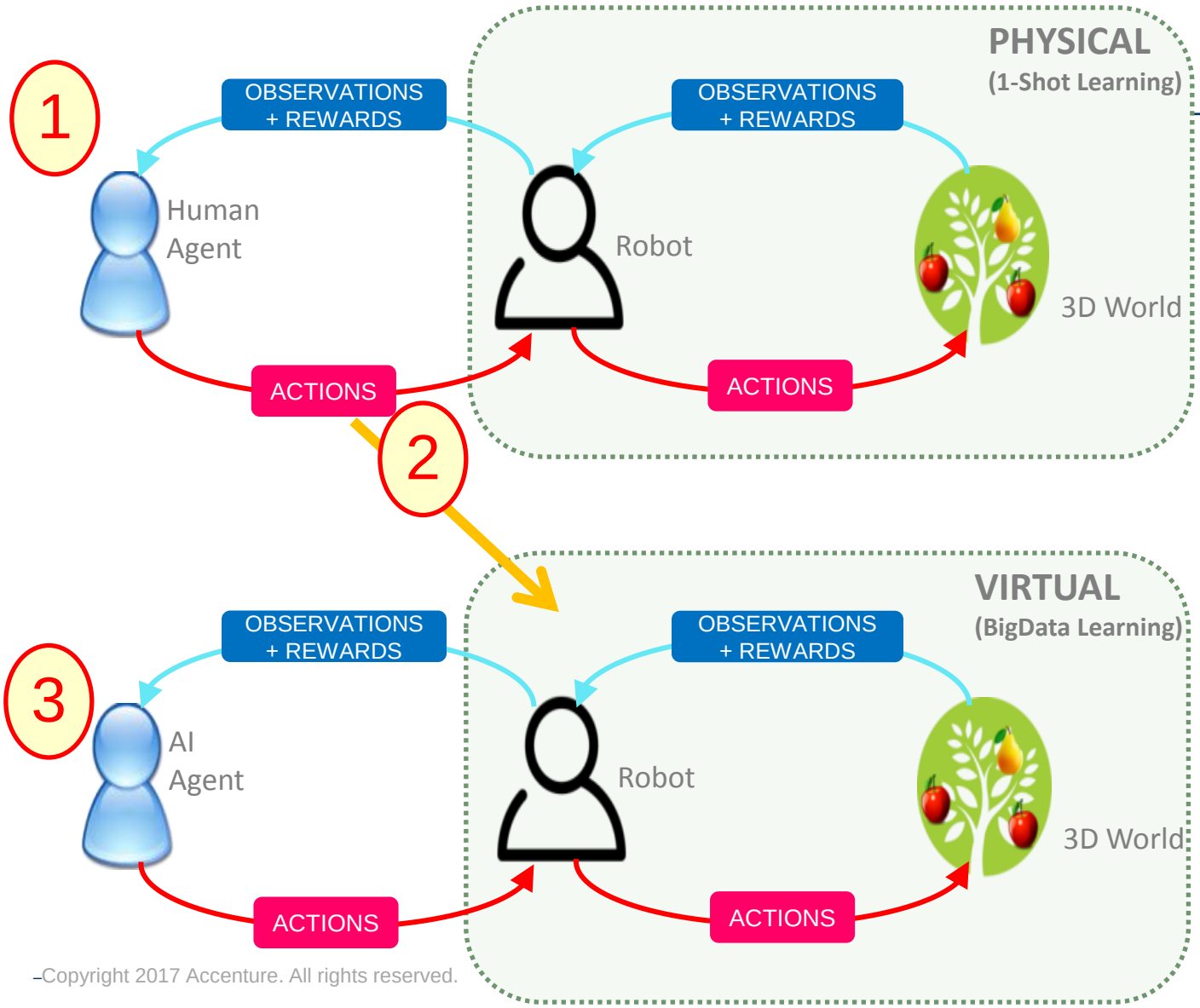
Apple Picker Roadmap



1. Human Agent controls a real Robot in real 3D world relying on video only
2. Human Agent controls a virtual Robot in virtual 3D world (video-game)

Copyright 2017 Accenture. All rights reserved.

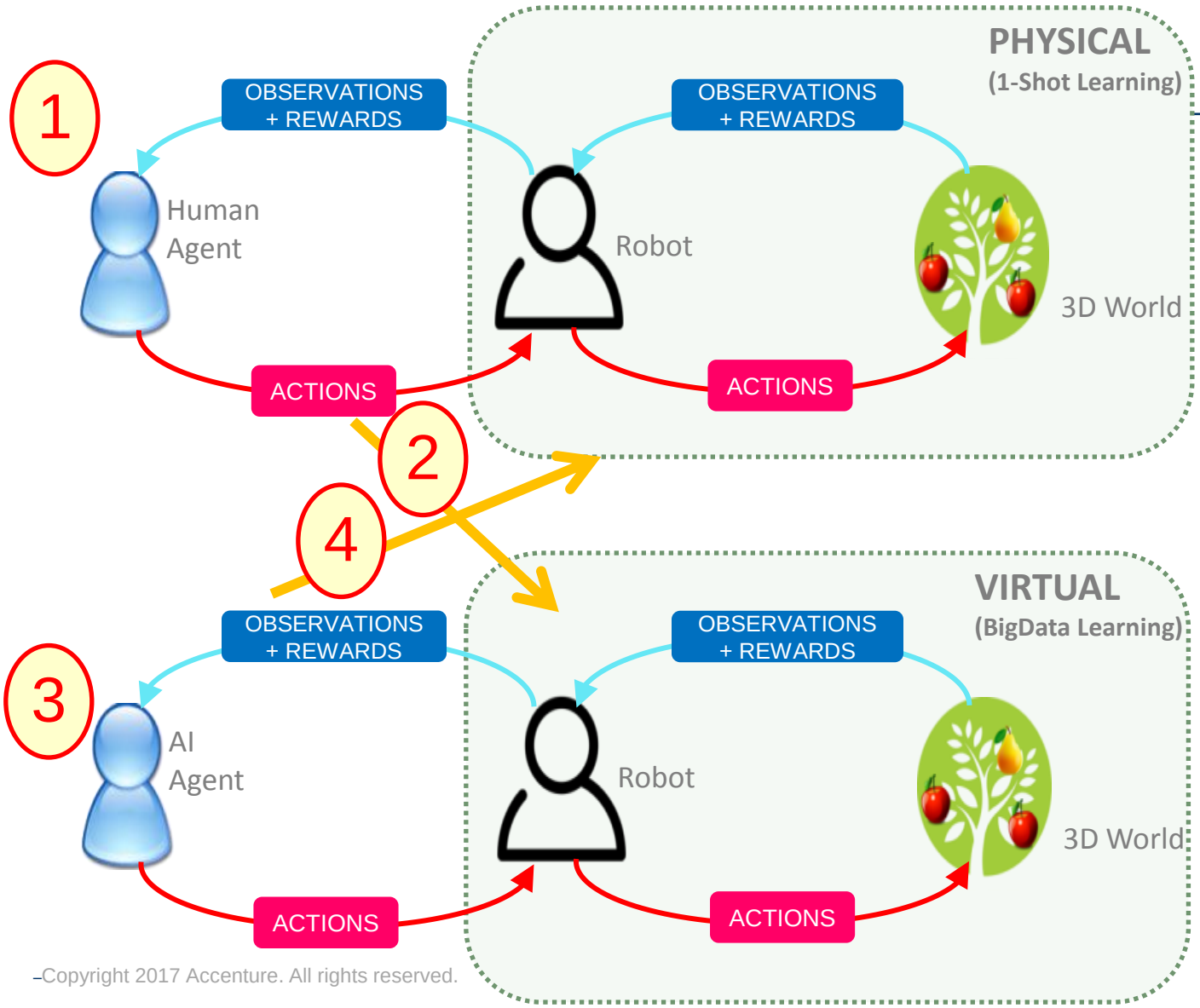
Apple Picker Roadmap



1. Human Agent controls a real Robot in real 3D world relying on video only
2. Human Agent controls a virtual Robot in virtual 3D world (video-game)
3. Deep Reinforcement Learning used to train AI Agent to control a virtual Robot (through millions of failures)

–Copyright 2017 Accenture. All rights reserved.

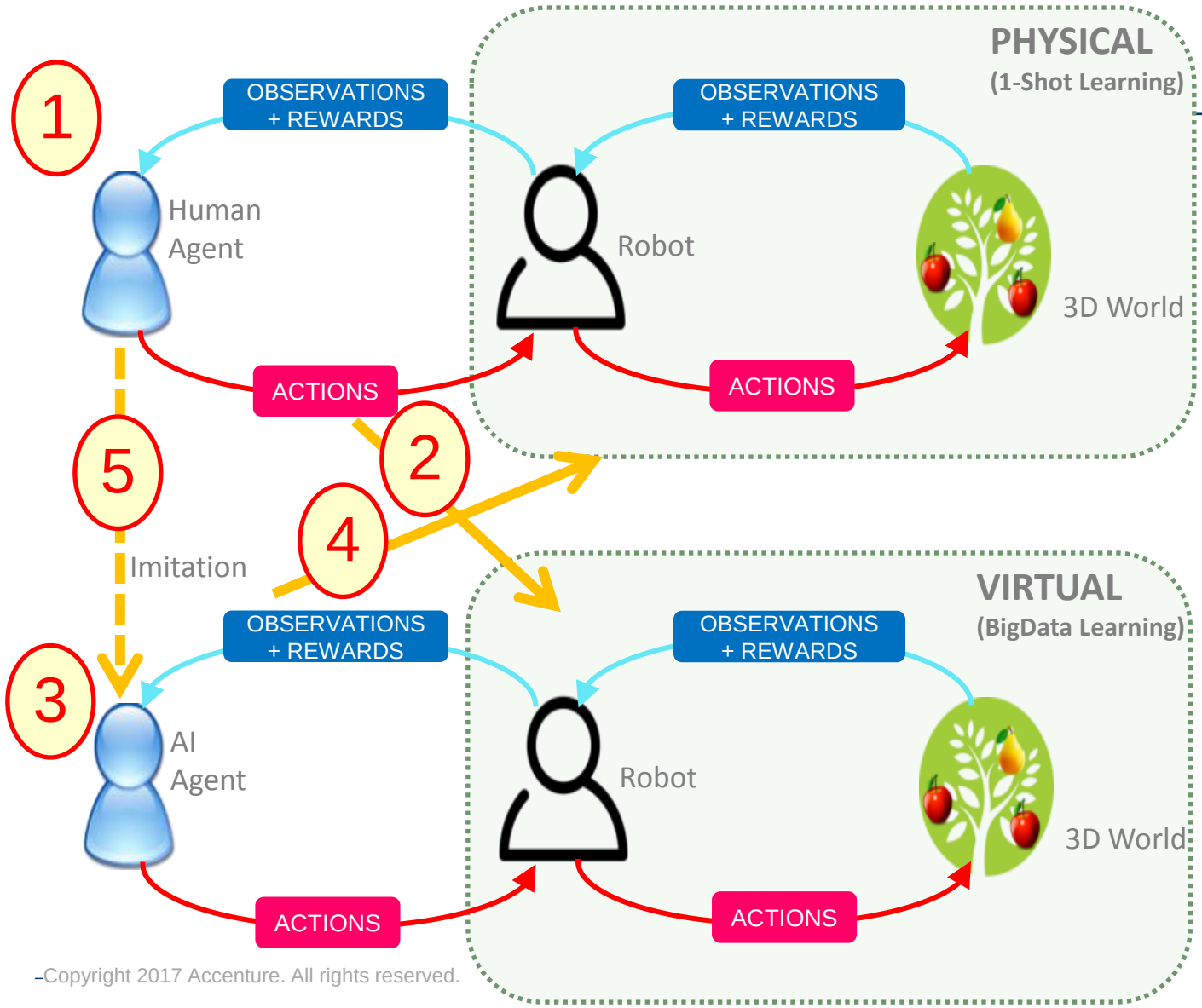
Apple Picker Roadmap



1. Human Agent controls a real Robot in real 3D world relying on video only
2. Human Agent controls a virtual Robot in virtual 3D world (video-game)
3. Deep Reinforcement Learning used to train AI Agent to control a virtual Robot (through millions of failures)
4. An AI Agent trained on large variety of virtual 3D worlds controls a real Robot in real 3D world

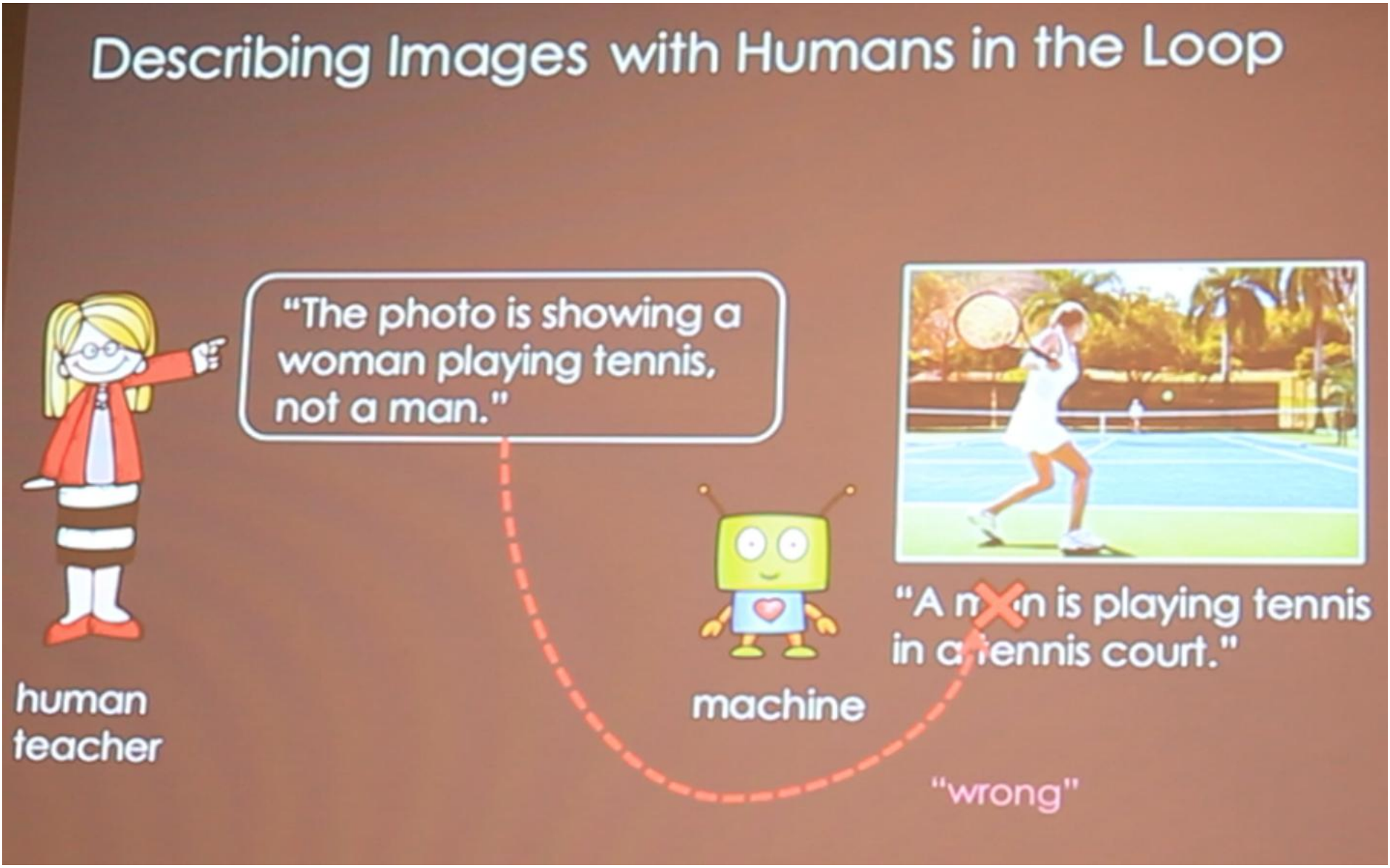
–Copyright 2017 Accenture. All rights reserved.

Apple Picker Roadmap



1. Human Agent controls a real Robot in real 3D world relying on video only
2. Human Agent controls a virtual Robot in virtual 3D world (video-game)
3. Deep Reinforcement Learning used to train AI Agent to control a virtual Robot (through millions of failures)
4. An AI Agent trained on large variety of virtual 3D worlds controls a real Robot in real 3D world
5. Improve the AI Agent performance by learning to imitate Human Agent actions

–Copyright 2017 Accenture. All rights reserved.



Connecting high-level semantics with low-level vision. Sanja Fidler

Describing Images with Humans in the Loop

human teacher

The UD taggingging is wrong, it should be Marisa DEPREL Butnere, because Gname always DEPREF Fname.

text = Šajā mācību gadā Aizkraukles novada ģimnāzijas 8. klasē mācījās Marisa Butnere no Amerikas.

ID	FORM	LEMMA	UPOSTAG	XPOSTAG	FEATS	HEAD	DEPREL	DEPS	MISC
1	Šajā	šī	DET	pd0fsln	Case=Loc|Gender=Fem|Number=Sing|PronType=Dem	3	det	3:det	
2	mācību	mācība	NOUN	ncfpg4	Case=Gen|Gender=Fem|Number=Plur	3	nmod	3:nmod:gen	
3	gadā	gads	NOUN	ncmsl1	Case=Loc|Gender=Masc|Number=Sing	9	obl	9:obl:loc	
4	Aizkraukles	Aizkraukle	PROPN	npfsg5	Case=Gen|Gender=Fem|Number=Sing	5	nmod	5:nmod:gen	
5	novada	novads	NOUN	ncmsg1	Case=Gen|Gender=Masc|Number=Sing	6	nmod	6:nmod:gen	
6	ģimnāzijas	ģimnāzija	NOUN	ncfsg4	Case=Gen|Gender=Fem|Number=Sing	8	nmod	8:nmod:gen	
7	8.	8.	ADJ	xo	NumType=Ord	8	amod	8:amod	
8	klasē	klase	NOUN	ncfsl5	Case=Loc|Gender=Fem|Number=Sing	9	obl	9:obl:loc	
9	mācījās	mācīties	VERB	vmyisi330an	Evident=Fh|Mood=Ind|Person=3|Polarity=Pos|Reflex=Yes	0	root	0:root	
10	Marisa	Marisa	PROPN	npfsn4	Case=Nom|Gender=Fem|Number=Sing	9	nsubj	9:nsubj	
11	Butnere	Butnere	PROPN	npfsn5	Case=Nom|Gender=Fem|Number=Sing	10	flat:name	10:flat:name	
12	no	no	ADP	spsg	_	13	case	13:case	
13	Amerikas	Amerika	PROPN	npfsg4	Case=Gen|Gender=Fem|Number=Sing	10	nmod	10:nmod:no	SpaceAfter=No
14	.	.	PUNCT	zs	_	9	punct	9:punct	

machine

A correct Universal Depencies labeling of the Latvian sentence «Šajā mācību gadā Aizkraukles novada ģimnāzijas 8. klasē mācījās Marisa Butnere no Amerikas.»

"wrong"

Describing Images with Humans in the Loop

human teacher

Kā tu to zini?!

machine

Marisa Butnere

Because of Visually grounded language learning!
AutoEncoder, HD-computing /Embedding, EpisodicMemory, KNN

Describing Images with Humans in the Loop

human teacher

Right!
Label this senetence with UD!

Meitenes vecvecāki līdz Otrajam pasaules karam dzīvojuši Latvijā, un arī viņa sevi uzskata par latvieti.

machine

Latvian text «Meitenes vecvecāki līdz Otrajam pasaules karam dzīvojuši Latvijā, un arī viņa sevi uzskata par latvieti.»

Describing Images with Humans in the Loop

human teacher

Great!

Meitenes vecvecāki līdz Otrajam pasaules karam dzīvojuši Latvijā, un arī viņa sevi uzskata par latvieti.

machine

ID	FORM	DEPS
1	Meitenes	2:nmod:gen
2	vecvecāki	7:nsubj
3	līdz	6:case
4	Otrajam	6:amod
5	pasaules	6:nmod:gen
6	karam	7:obl:līdz
7	dzīvojuši	0:root
8	Latvijā	7:obl:loc
9	,	14:punct
10	un	14:cc
11	arī	12:discourse
12	viņa	14:nsubj
13	sevi	14:obj
14	uzskata	7:conj
15	par	16:case
16	latvieti	14:xcomp
17	.	7:punct

Varu ieteikt studēt Latvijā
Apskatīt komentārus (0)

20.04.2011

Attēlā: Aristoteļa svētkos 2009. gadā. Ģirts Mandrovicks, Danute Ražuka-Ēbela, Marisa Butnere un Laura Sviklāne

Varu ieteikt studēt Latvijā

Agates Vucinas saruna ar Amerikas latvieti Marisu Butneri, Latvijas Universitātes medicīnas studenti

Kāpēc tavi vecāki atbalstīja tavu vēlmi dzīvot un mācīties Latvijā?

Mani vecāki vienmēr atbalstījuši savus bērnus. Brālis Dainis gribēja mācīties Vestpointas Militārajā akadēmijā, viņi atbalstīja arī mana brāļa Kristapa studiju nodomus. Es izvēlējos studēt medicīnu Rīgā. Man bija 18 gadu, kad atbraucu uz Latviju mācīties, taču es runāju ar citiem latviešiem Amerikā, kas man jautāja: – Kāpēc tieši uz Latviju, kāpēc nepaliki Amerikā? Man patīk Amerika, taču man ļoti patīk arī Latvija.

Daudzi Amerikas latvieši Latvijā pavada atvaļinājumu, bet mācās ASV augstskolās. Kādi ir tavi galvenie iemesli, kāpēc atbrauci studēt uz Latviju?

Amerikā mēs mācāmies latviešu skolās par Latviju, kultūru un tradīcijām, taču, kad pirmo reizi 13 gadu vecumā tēvs mani atveda brīvdienās uz Latviju, es iemīlēju šo zemi. Šeit ir grūti, taču visur pasaulē ir gan grūtības, gan arī labas īpašības. Šeit jūtos kā mājās, savā tēvzemē... Arī Amerikā tā jūtos, bet man šeit ir ļoti daudz radinieku. Amerikā man ir tikai vecāki, brāļi, vecmāmiņa un divas tantes, visi pārējie radi dzīvo šeit.

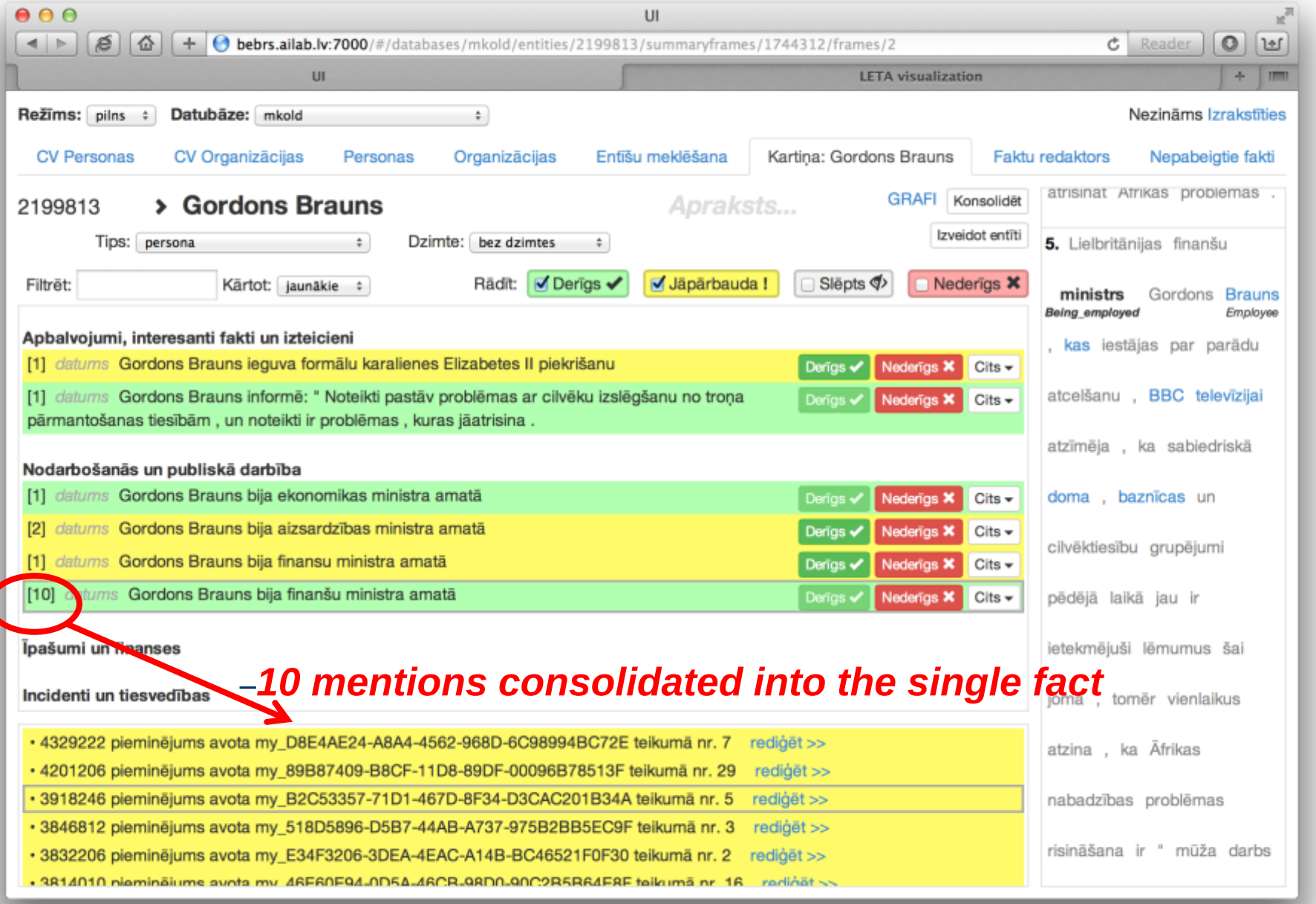
Amerikā viss notiek ļoti ātri, bet šeit cilvēki vairāk bauda dzīvi. Piemēram, universitātē man ir daudz jāmācās, bet es varu sarunāt ar pasniedzējiem, konkrētu laiku, kad likšu eksāmenus, kad rakstīšu kontroldarbus, ir labas attiecības starp pasniedzējiem, un visu var paspēt izdarīt. Kopš 2009. gada septembra darbojos LU Medicīnas fakultātes studentu pašpārvaldē.

http://www.laiks.us/jauno-laiks?mobile=ok&gads=2011&b=1&lbu=14050

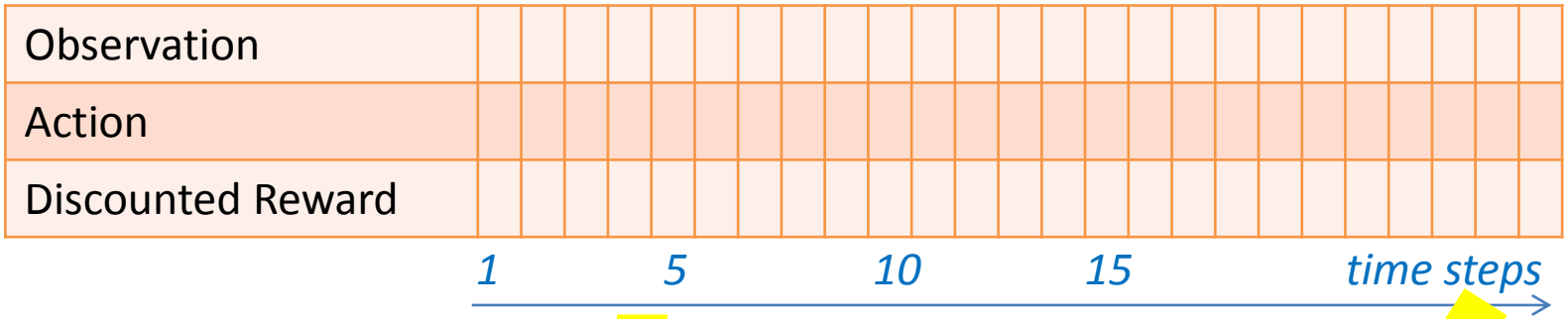
Potenciāls LZP granta pieteikums

- Potenciāli NLP beyond gisting (accuracy)
- Spiežam uz «Human in the Loop AI»
- Spiežam uz «Visually Grounded Language»
- Spiežam uz Sim2Real robotu
- Ar LETA CV sistēmu mes tur jau gandrīz bijām
  - Interfeisā varēja labot AI kļūdas
  - Tikai nepratām tos labojumus ņemt vērā

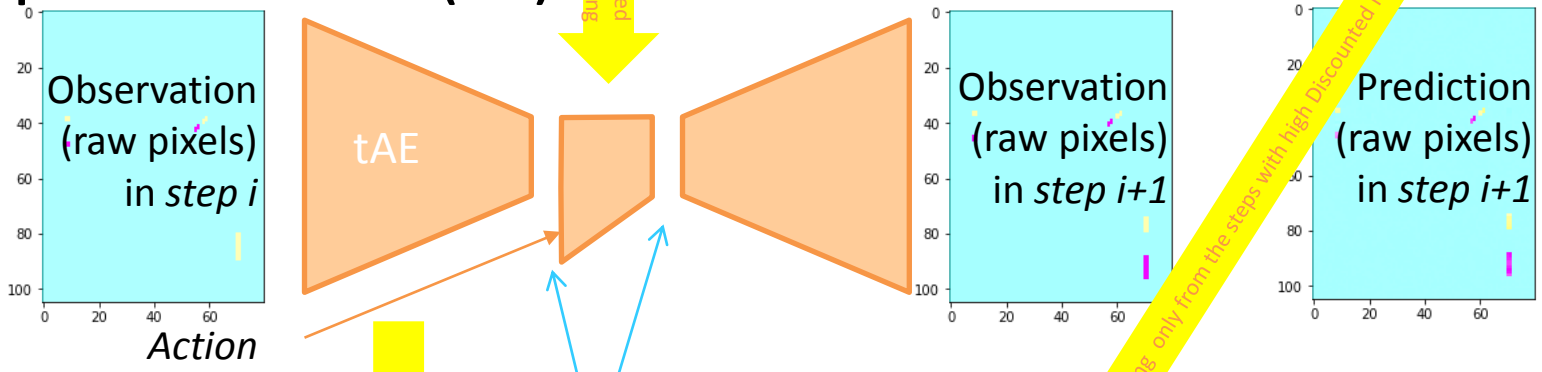
Profile Extractor (LETA)



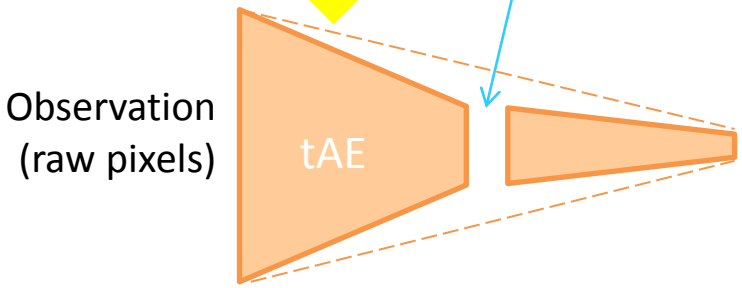
Episodic Memory



temporal AutoEncoder (tAE)



Policy Gradient RL



Episodic Control

- no forgetting (like DB)
- single neuron per time step
- in FPGA implementation discounted reward lookup time for kNN not dependent on EM size (like RAM)

https://github.com/LUMII-AILab/dBaby (runs in docker pull alvitunk/datz6056)

Top jauns H2020 pieteikums

- Būtībā SUMMA-2: beyond gisting (accuracy)
- Spiežam uz «Human in the Loop AI»
- Ar LETA CV sistēmu mes tur jau gandrīz bijām
  - Interfeisā varēja labot AI kļūdas
  - Tikai nepratām tos labojumus ņemt vērā

Grounding in visual environment

Statistical language models learned from text-only corpuses form the dominant paradigm in modern natural language understanding. While effective for many applications, these text-only distributional approaches suffer from limited semantics as they miss the interactive environment in which communication often takes place, i.e., its symbols are not grounded. Statistical approach also ignores interactive feedback from the end-user (correcting the errors produced by the learned model) due to statistical insignificance of the single corrected error.

The alternative “Human in the loop AI” paradigm is guided by the Reinforcement Learning via grounded one-shot learning from the end-user feedback (often regarded as Teacher).

Grounded one-shot learning is akin to data input into database – each new or modified fact there needs to be entered only once thanks to unambiguous identification (grounding / linking) of the involved entities. The legacy approach in natural language understanding has been linking selected entities to Wikipedia articles, but recently visual grounding of language has demonstrated some success in the Visually-Grounded Interaction and Language (ViGIL) NIPS 2017 Workshop (https://nips2017vigil.github.io/). Language grounding into visual models enables Reinforcement Learning to be applied to the natural language understanding problems.

Deep learning approaches are promising for grounding because they are capable of learning high-level semantics from low-level sensory data in both computer vision and language: embeddings learned from text modality are shown to enable zero-shot learning in visual modality and vice versa via descriptions, e.g. “zebra is like a horse only with stripes”. In addition, deep reinforcement learning provides an elegant framework for grounded language acquisition from interaction with an external environment or a Teacher.