

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

TESLA, INC.,
Petitioner,

v.

AUTONOMOUS DEVICES, LLC
Patent Owner

**PETITION FOR *INTER PARTES* REVIEW
OF U.S. PATENT NO. 11,238,344**

TABLE OF CONTENTS

	Page
I. MANDATORY NOTICES	1
A. Real Party-in-Interest	1
B. Related Matters.....	1
C. Lead and Back-up Counsel and Service Information	1
II. GROUNDS FOR STANDING.....	2
III. INTRODUCTION	2
IV. STATUTORY GROUNDS FOR THE CHALLENGES	4
V. LEVEL OF ORDINARY SKILL IN THE ART	4
VI. CLAIM CONSTRUCTION	4
A. “means for processing” (claim 20).....	5
VII. OVERVIEW OF THE '344 PATENT	5
VIII. OVERVIEW OF THE PRIOR ART	8
A. Grotmol.....	8
B. Hickman	9
IX. IDENTIFICATION OF HOW THE CHALLENGED CLAIMS ARE UNPATENTABLE.....	11
A. Ground 1	11
1. Claim 1	11
a. [1.1] A system comprising:	11
b. [1.2] one or more processors configured to perform at least:	11

c.	[1.3] accessing a memory that stores at least a knowledgebase that includes: a first correlation including a first circumstance representation correlated with a first one or more instruction sets for operating a first device, wherein the first circumstance representation represents a first circumstance detected at least in part by one or more sensors of the first device;	12
d.	[1.4] and wherein at least a portion of the first circumstance representation or at least a portion of the first one or more instruction sets for operating the first device is learned in a learning process that includes operating the first device at least partially by a user;.....	16
e.	[1.5] generating or receiving a second circumstance representation, wherein the second circumstance representation represents a second circumstance detected at least in part by: the one or more sensors of the first device or at least in part by one or more sensors of a second device;.....	17
f.	[1.6] anticipating the first one or more instruction sets for operating the first device learned in the learning process based on at least partial match between the second circumstance representation and the first circumstance representation;	18
g.	[1.7] and at least in response to the anticipating, executing the first one or more instruction sets for operating the first device learned in the learning process, wherein the first device or the second device autonomously performs one or more operations defined by the first one or more instruction sets for operating the first device learned in the learning process.	20
2.	Claim 2	22
3.	Claim 3	23

4.	Claim 4	25
5.	Claim 5	26
6.	Claim 7	27
a.	[7.1] The system of claim 3, wherein the one or more processors are further configured to perform at least: modifying: the first one or more instruction sets for operating the first device, or a copy of the first one or more instruction sets for operating the first device,	27
b.	[7.2] and wherein the anticipating the first one or more instruction sets for operating the first device based on the at least partial match between the second circumstance representation and the first circumstance representation includes: anticipating the modified the first one or more instruction sets for operating the first device based on the at least partial match between the second circumstance representation and the first circumstance representation, or anticipating the modified the copy of the first one or more instruction sets for operating the first device based on the at least partial match between the second circumstance representation and the first circumstance representation, and.....	28
c.	[7.3] wherein the executing the first one or more instruction sets for operating the first device includes: executing the modified the first one or more instruction sets for operating the first device, or executing the modified the copy of the first one or more instruction sets for operating the first device, and wherein the autonomously performing, by the first device or by the second device, the one or more operations defined by the first one or more instruction sets for operating the first device includes: autonomously performing, by the first	

device or by the second device, one or more operations defined by the modified the first one or more instruction sets for operating the first device, or autonomously performing, by the first device or by the second device, one or more operations defined by the modified the copy of the first one or more instruction sets for operating the first device.29

7. Claim 8:30

a. [8.1] The system of claim 3, wherein the one or more processors are further configured to perform at least: modifying at least one of: the first circumstance representation, a copy of the first circumstance representation, the second circumstance representation, or a copy of the second circumstance representation,30

b. [8.2] and wherein the at least partial match between the second circumstance representation and the first circumstance representation includes: (i) at least partial match between the modified the second circumstance representation and the first circumstance representation, ... (iii) at least partial match between the second circumstance representation and the modified the first circumstance representation, ... [or] (v) at least partial match between the modified the second circumstance representation and the modified the first circumstance representation,31

8. Claim 932

9. Claim 1333

10. Claim 1435

a. [14.1] The system of claim 13, wherein the artificial intelligence system includes: one or more inputs for receiving one or more object representations, and

	one or more outputs for providing one or more instruction sets,	35
b.	[14.2] and wherein the learning, at least in part by the artificial intelligence system, the at least the portion of the first circumstance representation or the at least the portion of the first one or more instruction sets for operating the first device includes applying the first one or more object representations to the one or more inputs of the artificial intelligence system,	36
c.	[14.3] and wherein the anticipating, at least in part by the artificial intelligence system, the first one or more instruction sets for operating the first device includes: applying the second one or more object representations to the one or more inputs of the artificial intelligence system; and receiving the first one or more instruction sets for operating the first device from the one or more outputs of the artificial intelligence system.	37
11.	Claim 15	37
12.	Claim 16	38
13.	Claim 17	39
14.	Claim 18	40
a.	[18.1] The system of claim 3, wherein at least some elements of the system are included in: a single device, or multiple devices, and	40
a.	[18.2] wherein the one or more processors include: one or more microcontrollers, one or more computing devices, or one or more electronic devices, and.....	40
b.	[18.3] wherein the memory includes: a non-transitory machine readable medium, a volatile	

- memory, a non-volatile memory, a storage device,
or a storage system, and.....40
- c. [18.4] wherein the one or more sensors of the first
device include at least one of: one or more
cameras... and wherein the one or more sensors of
the second device include at least one of: one or
more cameras, ...and41
- d. [18.5] wherein the first device includes: a robot, a
vehicle, ... or a mechanical machine, and.....41
- e. [18.6] wherein the second device includes: a robot,
a vehicle, ... or a mechanical machine, and41
- f. [18.7] wherein the first circumstance representation
is a data structure that includes one or more data
about the first circumstance, and wherein the second
circumstance representation is a data structure that
includes one or more data about the second
circumstance,42
- g. [18.8] and wherein the first circumstance includes
one or more objects detected ... by the one or more
sensors of the first device ... during a first time
period, and wherein the second circumstance
includes: one or more objects detected ... by the one
or more sensors of the first device ... during a
second time period, or one or more objects detected
... by the one or more sensors of the second device
... during a second time period, and.....44
- h. [18.9] wherein an instruction set of the first one or
more instruction sets for operating the first device
includes at least one of: only one instruction,
multiple instructions, ... one or more commands, ...
or one or more information, and.....46
- i. [18.10] wherein the first one or more object
representations include: one object representation,
... and wherein the second one or more object

	representations include: one object representation, ... and wherein the first one or more object representations include: ... one or more digital pictures that depict one or more objects, ... and wherein the second one or more object representations include: ... one or more digital pictures that depict one or more objects, ... and wherein the at least the portion of the first circumstance representation includes: at least a portion of one object representation ... and	46
j.	[18.11] wherein the at least the portion of the first one or more instruction sets for operating the first device includes: at least a portion of one instruction set of the first one or more instruction sets for operating the first device	47
k.	[18.12] and wherein: an object of the first circumstance is the same as an object of the second circumstance, multiple objects of the first circumstance are the same as multiple objects of the second circumstance, all objects of the first circumstance are the same as all objects of the second circumstance, or all objects of the first circumstance are different than all objects of the second circumstance.	47
15.	Claim 19	48
16.	Claim 20	50
B.	Ground 2	53
1.	Claim 10	53
2.	Claim 11	54
C.	Ground 3	55
1.	Motivation to Combine Grotmol and Hickman	55

2.	Claim 6	58
3.	Claim 11	61
4.	Claim 12	63
a.	[12.1] The system of claim 3, wherein the knowledgebase further includes a third circumstance representation correlated with a second one or more instruction sets for operating a third device, and wherein the third circumstance representation represents a third circumstance detected at least in part by one or more sensors of the third device, and wherein the one or more processors are further configured to perform at least:	63
b.	[12.2] generating or receiving a fourth circumstance representation, wherein the fourth circumstance representation represents a fourth circumstance detected at least in part by one or more sensors of a fourth device; anticipating the second one or more instruction sets for operating the third device based on at least partial match between the fourth circumstance representation and the third circumstance representation; and at least in response to the anticipating the second one or more instruction sets for operating the third device, executing the second one or more instruction sets for operating the third device, wherein the fourth device autonomously performs one or more operations defined by the second one or more instruction sets for operating the third device.	64
X.	CONCLUSION.....	65

PETITIONER’S EXHIBIT LIST

Ex. 1001	U.S. Patent No. 11,238,344 (the “’344 patent”)
Ex. 1002	Declaration of Jason Janet, Ph.D.
Ex. 1003	Curriculum Vitae of Jason Janet, Ph.D.
Ex. 1004	U.S. Patent No. 9,604,359 (“Grotmol”)
Ex. 1005	U.S. Patent No. 8,639,644 (“Hickman”)
Ex. 1006	Service, Robert F., <i>Minds of Their Own</i> , Sci. Magazine. Oct. 10, 2014, 346(6206):182-183, available at, https://www.usna.edu/AcResearch/_files/documents/Science_Minds_of_Their_own_182.full1.pdf
Ex. 1007	Prosecution File History of U.S. Patent No. 11,238,344

I. MANDATORY NOTICES

A. Real Party-in-Interest

Tesla, Inc. is the sole real party-in-interest.

B. Related Matters

The '344 patent is the subject of a civil action, *Autonomous Devices LLC v. Tesla, Inc.*, No. 22-cv-1466 (MN) (D. Del.), filed November 7, 2022.

The '344 patent is a continuation of U.S. Patent Application No. 15/340,991, filed on November 2, 2016, now U.S. Patent No. 10,452,974 (“the '974 patent”).

Tesla has also filed a petition for *inter partes* review of the '974 patent.

C. Lead and Back-up Counsel and Service Information

Anish Desai (lead counsel)
Reg. No. 73,760
Ian Moore (backup counsel)
Reg. No. 77,395
Weil, Gotshal & Manges LLP
767 Fifth Ave.
New York, NY 10153
T: 212-310-8000
F: 212-310-8007
anish.desai@weil.com
ian.moore@weil.com

Adrian Percer (backup counsel)
Reg. No. 46,986
Weil, Gotshal & Manges LLP
201 Redwood Shores Parkway
Redwood Shores, CA 94065
T: 650-802-3000
F: 650-802-3100
adrian.percer@weil.com

Christopher Pepe (backup counsel)
Reg. No. 73,851
Matthew Sieger (backup counsel)
Reg. No. 76,051
Taylor J. Lawrence (backup counsel)
Reg. No. 79,817
Allison D. Herzig (backup counsel)
Reg. No. 80,747
Weil, Gotshal & Manges LLP
2001 M Street NW, Suite 600
Washington, DC 20036
T: 202-682-7000
F: 202-857-0940
christopher.pepe@weil.com
matthew.sieger@weil.com
taylor.lawrence@weil.com
allison.herzig@weil.com

II. GROUNDS FOR STANDING

Petitioner certifies that the '344 patent is available for *inter partes* review and that Petitioner is not barred or estopped from requesting *inter partes* review challenging the patent claims on the grounds identified in this Petition.

III. INTRODUCTION

The '344 patent claims a system for autonomous device operation in which: a memory stores circumstance representations (*e.g.*, pictures of the device's circumstances) correlated with instructions for operating a device where the instructions or circumstance representations are learned during operation of the device by a user; new circumstance representations are generated or received; instructions for operating the device are anticipated by determining a partial match between the new circumstance representations and the stored circumstance representations; and the correlated instructions are executed to autonomously operate the device. As will be demonstrated in this Petition, the prior art anticipates and/or renders obvious claims 1-20 of the '344 patent.

Petitioner requests *inter partes* review based upon U.S. Patent No. 9,604,359 ("Grotmol") and U.S. Patent No. 8,639,644 ("Hickman"), neither of which were before the Examiner during prosecution. Grotmol discloses the same approach to autonomous operation claimed in the '344 patent:

A robotic vehicle may utilize a camera and stores the sequence of images of a visual scene seen when following a trajectory

during training in an ordered buffer. Motor commands associated with a given image may be stored. During autonomous operation, an acquired image may be compared with one or more images from the training buffer in order to determine the most likely match.

Ex. 1004, Abstract. Grotmol further teaches that a robot's "trained configuration may be stored" and "loaded to one or more other robots in order to provide learned behaviors." Ex. 1004, 33:12-35.

Hickman—a reference cited in Grotmol—further discloses "shared robot knowledge bases for use with cloud computing systems," in which the cloud computing system "collects data from a robot about an object the robot has encountered in its environment, and stores the received data in the shared robot knowledge base . . . sends instructions for interacting with an object to a robot . . . [and] sends instructions to a robot for executing an application based on information stored in the shared robot knowledge base." Ex. 1005, Abstract.

Grotmol anticipates and/or renders obvious in combination with Hickman claims 1-20 of the '344 patent. Petitioner respectfully requests that the Board institute review.

IV. STATUTORY GROUNDS FOR THE CHALLENGES

Ground 1	Anticipation of claims 1-5, 7-9, and 13-20 based on Grotmol.
Ground 2	Obviousness of claims 10 and 11 based on Grotmol in view of the knowledge of a POSITA.
Ground 3	Obviousness of Claims 6, 11, and 12 based on Grotmol in view of Hickman.

V. LEVEL OF ORDINARY SKILL IN THE ART

A person having ordinary skill in the art (“POSITA”) relevant to the ’344 patent as of November 2, 2016—the earliest possible priority date of the ’344 patent—would have at least: (1) a bachelor’s degree in electrical engineering, computer engineering, computer science, systems engineering, or equivalent course work with two to three years of work experience in computer vision, image/video processing, and/or machine learning; or (2) a master’s degree in electrical engineering, computer engineering, computer science, systems engineering, or equivalent course work with a focus in computer vision, image/video processing, and/or machine learning. Ex. 1002 (“Decl.”), ¶¶ 40-42.

VI. CLAIM CONSTRUCTION

The Petition applies the ordinary and customary meaning of the claim terms. Where appropriate, Petitioner explains the meaning of claim terms in its analysis of

how the prior art anticipates and/or renders obvious the challenged claims. The following section addresses the construction of a term under 35 U.S.C. § 112(f).

A. “means for processing” (claim 20)

Claim 20 requires a “means for processing” that performs the following functions: “generating or receiving a second circumstance representation ...,” “anticipating the first one or more instruction sets for operating the first device ...,” and “executing the first one or more instruction sets ...” Ex. 1001, claim 20. The ’344 patent specification describes that the corresponding structure for the claimed functions comprises a processor or processor circuit. *Id.* at 17:9-42 (stating that the system includes a “processor circuit” that performs operations of receiving circumstance representations, receiving a new digital picture, anticipating the instruction set, and executing the instruction set); claim 1 (requiring a processor that performs the receiving, anticipating, and executing steps); 23:46-54, 58:64-59:41. Accordingly, a POSITA would understand that the corresponding structure for the means for processing comprises a processor circuit or processor. *See also* Decl. ¶ 58.

VII. OVERVIEW OF THE ’344 PATENT

The ’344 patent was filed on August 14, 2019, and is a continuation of Application No. 15/340,991, filed on November 2, 2016, issued as U.S. Patent No. 10,452,974.

The disclosure of the '344 patent relates to “devices, systems, methods, and/or applications for learning a device's operation in various circumstances, storing this knowledge in a knowledgebase (i.e. neural network, graph, sequences, etc.), and enabling autonomous operation of the device.” Ex. 1001, Abstract. More specifically, the '344 patent describes a method of automation comprising “learning a device’s circumstances including objects with various properties along with correlated instruction sets for operating the device, storing this knowledge in a knowledgebase (i.e. neural network, graph, sequences, etc.), and operating a device autonomously.” Ex. 1001, 57:58-58:5. The '344 patent refers to this approach as “device’s circumstances for autonomous device operation,” or “DCADO.” *Id.* Figure 32, copied below, illustrates this functionality.

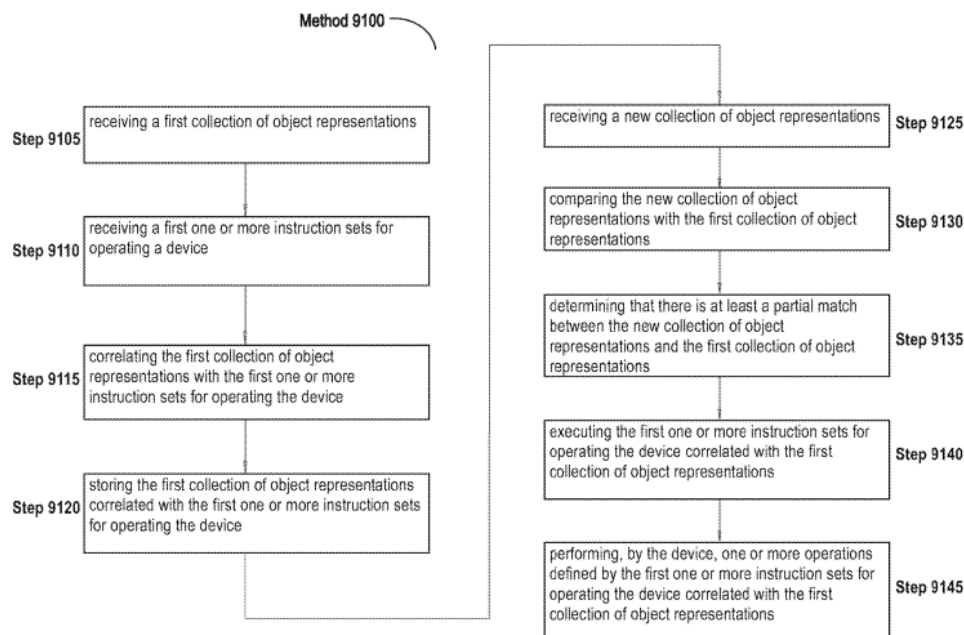


FIG. 32

Ex. 1001, Fig. 32

The '344 patent further describes that the DCADO functionality can be implemented on a “Server 96 accessible over Network 95,” such that “[a]ny number of Devices 98 may connect to such remote DCADO Unit 100 and the remote DCADO Unit 100 may learn their operations in circumstances including objects with various properties. In turn, any number of Devices 98 can utilize the remote DCADO Unit 100 for autonomous operation in circumstances including objects with various properties.” Ex. 1001, 95:21-29.

During prosecution of the '344 patent, the claims were allowed over Schnittman (U.S. 2016/0167226). Ex. 1007. The inventor submitted a statement after allowance that the “training in Schnittman is different than the process recited in [the claims] because training in Schnittman does not include a user at least partially operating a device...and what is learned in the training in Schnittman is different than what is learned in the learning process recited by [the claims].” *Id.* (12/26/2021 Comments on Statement of Reasons for Allowance) at 2. As detailed below, Grotmol discloses training in which a user operates a robotic device, and what is learned during the training includes images (circumstance representations) correlated with motor commands (instruction sets) as claimed in the '344 patent.

VIII. OVERVIEW OF THE PRIOR ART

A. Grotmol

Grotmol is a U.S. patent filed on January 27, 2015, which issued on March 28, 2017. Grotmol is prior art under at least 35 U.S.C. §§ 102(a) and 102(d).

Grotmol relates to “[a]n apparatus and methods for training and/or operating a robotic device to follow a trajectory.” Ex. 1004, Abstract. Grotmol describes that “conventional approaches” to training a robot entailed a user “first control[ing] the robot one time or multiple times to move along the desired path.” Ex. 1004, 1:62-67. That would constitute the “training of the robot,” and thereafter, the robot would perform the same navigation autonomously. *Id.*, 1:67-2:2. A “typical approach” was to store the motor commands executed by a user during training then “simply replay them.” *Id.*, 2:3-5. Grotmol then describes its own approach, whereby user commands are associated with camera images captured during training, such that during “autonomous operation, an acquired image may be compared with one or more images from the training buffer” to determine how to control the robot. Ex. 1004, Abstract.

More specifically, Grotmol describes a “listen mode,” in which a robot observes a teacher’s commands in order to “learn an association.” Ex. 1004, 23:3-11. Grotmol also describes an “override mode,” in which a robot will operate autonomously until a teacher sends a command, in which case the robot will apply

the teacher's command and "try to remember it for the next time the same situation occurs." *Id.*, 23:12-22. "During training, images (e.g., raw and/or pre-processed) may be stored in a memory buffer (training buffer)," along with "[m]otor commands issued by a trainer corresponding to time instances when the images are acquired." *Id.*, 39:13-22. The robot may then switch—either manually or automatically—into an "autonomous mode," in which the robot performs tasks autonomously based on what it has learned. *Id.*, 23:46-48, 24:1-16. "During autonomous operation, control process of the robot may be configured to compare a given (e.g., the most recent, current) image with one or more the images from the training buffer." *Id.*, 39:23-28. "In some implementations, motor commands at a given time step may be obtained by taking the stored motor commands from the training buffer that may correspond to the best matching stored image." *Id.*, 43:5-9.

Grotmol further teaches that a robot's "trained configuration may be stored" and "loaded to one or more other robots in order to provide learned behaviors." Ex. 1004, 33:12-35.

B. Hickman

Hickman is a U.S. patent filed on May 4, 2012, which issued on January 28, 2014. Hickman is prior art under at least 35 U.S.C. §§ 102(a) and 102(d). Hickman is cited in Grotmol. Ex. 1004.002.

The disclosure of Hickman relates to “shared robot knowledge bases for use with cloud computing systems.” Ex. 1005, Abstract. Hickman teaches a method whereby “information in the shared robot knowledge bases is updated based on robot experiences so that any particular robot may benefit from prior experiences of other robots.” *Id.* One embodiment of Hickman’s shared robot knowledge base is illustrated in Figure 4, copied below. *See* Ex. 1005, 25:3-27:26 (describing figure).

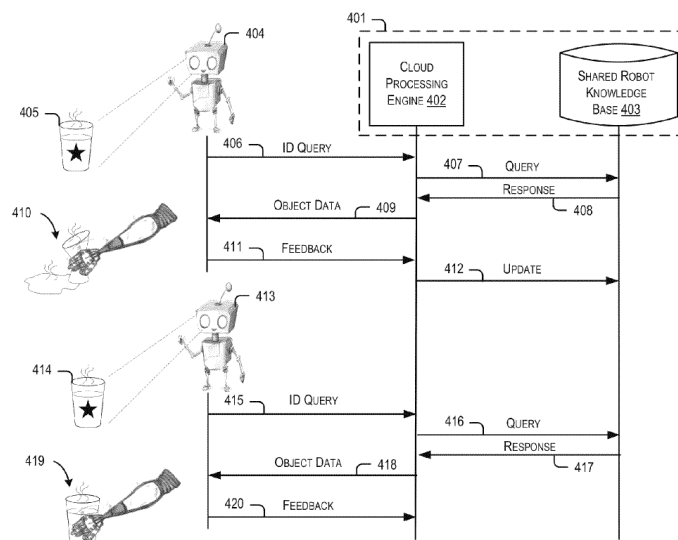


FIGURE 4
Ex. 1005, Fig. 4

As shown above, a first robot 404 encounters an object 405 and obtains image data associated with it. Ex. 1005, 25:14-26. The first robot attempts to grasp the object but crushes it (410). *Id.*, 26:38-56. “When an accident happens, a human can ‘coach’ the robot on how much force to use by, for example, manually controlling the robot’s hand.” *Id.* “The cloud processing system 402 can then update” with this new training. *Id.* Later, a second robot may encounter an object 414 with similar

image data to object 405. *Id.*, 26:57-27:6. The second robot may attempt to grasp the new object—but with less force based on the updated training data from the first robot—and this time be successful. *Id.*, 26:66-27:26.

IX. IDENTIFICATION OF HOW THE CHALLENGED CLAIMS ARE UNPATENTABLE

A. Ground 1

Grotmol anticipates claims 1-5, 7-9, and 13-20 of the '344 patent.

1. Claim 1

a. [1.1] A system comprising:

To the extent the preamble is limiting, Grotmol discloses a system implemented as a “self-contained robotic apparatus.” Ex. 1004, 3:27-30; *see also* Abstract (“An apparatus and methods for training and/or operating a robotic device to follow a trajectory.”); 10:45-48. Accordingly, Grotmol discloses a system. Decl. ¶ 163.

b. [1.2] one or more processors configured to perform at least:

Grotmol discloses limitation [1.2]. Decl. ¶¶ 164.

Grotmol illustrates the system including a “processing” unit in Figure 7 that is connected to, *inter alia*, a sensory unit and a memory unit. Ex. 1004, Fig. 7.

Grotmol further defines “processor” as follows:

As used herein, the terms “processor”, “microprocessor” and “digital processor” are meant generally to include all types of digital processing devices including, without limitation, digital

signal processors (DSPs), reduced instruction set computers (RISC), general-purpose (CISC) processors, microprocessors, gate arrays (e.g., field programmable gate arrays (FPGAs)), PLDs, reconfigurable computer fabrics (RCFs), array processors, secure microprocessors, and application-specific integrated circuits (ASICs). Such digital processors may be contained on a single unitary IC die, or distributed across multiple components.

Id., 8:5-15. Grotmol also discloses “a self-contained robotic apparatus, the apparatus comprising ... *one or more physical processors* configured to operate [the system].”¹

Id., claim 12; *see also* Fig. 8.

- c. **[1.3] accessing a memory that stores at least a knowledgebase that includes: a first correlation including a first circumstance representation correlated with a first one or more instruction sets for operating a first device, wherein the first circumstance representation represents a first circumstance detected at least in part by one or more sensors of the first device;**

Grotmol discloses limitation [1.3]. Decl. ¶ 165-172.

Grotmol discloses a “memory component may be configured to store training sets,” where “[a] given training set may comprise an instance of the video frame, an instance of the binary signal, and an instance of a motor control indication configured to cause the robot to execute an action.” Ex. 1004, 3:27-39; Abstract (“A robotic vehicle may utilize a camera and stores the sequence of images of a visual scene seen when following a trajectory during training in an ordered buffer. Motor

¹ All emphasis has been added, unless noted otherwise.

commands associated with a given image may be stored.”); *see also* 3:31-39, 8:66-9:3, 10:23-26, 38:36-58, 39:13-22. Thus, Grotmol discloses a memory that stores a knowledgebase² including multiple correlations (training sets), where each correlation includes a “circumstance representation”³ (image of scene from a robot vehicle’s operation) correlated with an “instruction set for operation of a first device”⁴ (motor control indication configured to cause the robot to execute an action).

² The ’344 patent states that a “[k]nowledgebase 530 may be or include **any data structure or arrangement** capable of storing the knowledge of a devices’s operation in circumstances including objects with various properties.” Ex. 1001, 102:15-18.

³ The ’344 patent specification does not define or even contain the term “circumstance representation.” The most similar term used in the specification, a “collection of object representations,” is described generally as “one or more representations of objects detected by the sensor at a time.” Ex. 1001, 3:42-44. The “sensor” that detects the object representations may include a “[c]amera ... for capturing one or more pictures” or a “motion picture camera that can capture streams of pictures (i.e. motion pictures, videos, etc.).” *Id.*, 69:52-70:32.

⁴ The ’344 patent describes that “an instruction set may be used or executed by a logic circuit ... [or] actuator in operating a device” Ex. 1001, 152:12-18. An instruction set can be “received from a processor, application program, logic circuit, and/or other processing element,” “from memory[], hard drive, or any other storage element or repository,” or “over a network such as Internet, local area network, wireless network, and/or other network.” *Id.*, 152:21-40. “In general, an instruction set can be received by any element of the system.” *Id.*, 152:44-45. “Instruction Set 526 may include one or more instructions or commands of a high-level programming language such as Java or SQL, a low-level language such as assembly or machine language, an intermediate language or construct such as bytecode, and/or **any other language or construct.**” *Id.*, 93:34-39.

The images in Grotmol are “associated with” or “correspond to”—*i.e.*, are correlated with⁵—motor control indications issued around the same time. Decl. ¶ 166. Grotmol describes training behaviors in a robotic device by “learning of *associations* between sensory context . . . and a respective action . . . during training.” Ex. 1004, 33:59-62; Abstract (“Motor commands *associated with* a given image may be stored.”). This association may be based on the timing of the images and the motor commands. Ex. 1004, 39:17-19 (“Motor commands issued by a trainer *corresponding to* time instances when the images are acquired may be stored.”); 43:5-9 (“[M]otor commands at a given time step may be obtained by taking the stored motor commands from the training buffer that may *correspond to* the best matching stored image.”).

Grotmol discloses that circumstance representations (images) represent circumstances detected by a sensor of the first device. Specifically, Grotmol discloses the video frame instances are detected by a “sensor component” of the robotic vehicle. Ex. 1004, 3:31-33 (“The first sensor component may be configured to provide a signal conveying a video frame comprising a plurality of pixels.”); Abstract (“A robotic vehical may utilize a camera and stores the sequence of images

⁵ The ’344 patent describes that a device may “correlate a Collection of Object Respresentations 525 with one or more ***temporally corresponding*** Instruction Sets . . . ”). Ex. 1001, 99:18-21.

of a visual scene seen when following a trajectory during training in an ordered buffer.”). As depicted in Figure 8 of Grotmol, the “memory” of Grotmol is connected to the “sensory” functionality through the “processing” functionality. *See* Ex. 1004, 9:1-3 (“[T]he output provided by the sensor component 166 may comprise a stream of pixel values associated with one or more digital images.”). Both Grotmol and the ’344 patent disclose that digital images may come from, for example, a CCD image sensor in the form of color encoded pixels. *Compare* Ex. 1001, 69:65-70:5, 112:26-32 *with* Ex. 1004, 9:8-22. Grotmol further discloses that these pixels are used for object (or circumstance) recognition. Ex. 1004, 8:66-9:3 (“In one or more implementations, such as object recognition, and/or obstacle avoidance, the output provided by the sensor component 166 may comprise a stream of pixel values associated with one or more digital images.”).

Grotmol further discloses accessing the training sets stored in the memory component to perform autonomous operation. Ex. 1004, Abstract (“During autonomous operation, an acquired image may be compared with one or more images from the training buffer in order to determine the most likely match.”); 39:23-28 (“During autonomous operation, control process of the robot may be configured to compare a given (e.g., the most recent, current) image with *one or more* the images from the training buffer.”).

Accordingly, Grotmol discloses accessing a memory that stores the claimed knowledgebase. Decl. ¶¶ 165-172.

- d. **[1.4] and wherein at least a portion of the first circumstance representation or at least a portion of the first one or more instruction sets for operating the first device is learned in a learning process that includes operating the first device at least partially by a user;**

Grotmol discloses limitation [1.4]. Decl. ¶ 173.

Grotmol discloses that a human (“user”⁶) can operate the robot during a training process and “[t]rained behaviors of the robotic device may be configured based on *learning* of associations between sensory context (e.g., presence of an obstacle in front of the robotic vehicle) and a respective action (e.g., right turn) during training.” Ex. 1004, 33:59-62 (emphasis added). In both the “listen mode” and “override mode,” Grotmol teaches that the robot will listen to commands sent by a trainer (user) “and try to remember it for the next time the same situation occurs.” Ex. 1004, 23:3-29; *see also id.*, 24:55-25:5 (“While training a particular behavior, the *user* may concentrate on a particular behavior and may not show any examples of other behaviors thus making a prediction task for the classifier substantially easier (e.g., less computationally intensive).”); 34:49-55 (“In some

⁶ “User 50 (also referred to simply as user or other similar name or reference) comprises a human user or non-human user.” Ex. 1001, 68:29-31.

implementations of supervised training of robots, control instructions (also referred to as corrections) produced by the *trainer* (e.g., human) may be combined with control instructions produced by the robot controller instructions (predictions).”). Grotmol describes this training as a learning process. *E.g., id.*, 24:17-20 (“Apparatus and methods for hierarchical *learning* are disclosed herein, in accordance with one or more implementations. In *supervised learning*, a user may train simple tasks by demonstrating example tasks to the system.”); claim 12 (reciting “one or more physical processors configured to operate a random k-nearest neighbors *learning process*”). Accordingly, Grotmol discloses a first circumstance representation and a first instruction set are learned during a learning process in which a user operates the first device. Decl. ¶ 173.

- e. **[1.5] generating or receiving a second circumstance representation, wherein the second circumstance representation represents a second circumstance detected at least in part by: the one or more sensors of the first device or at least in part by one or more sensors of a second device;**

Grotmol discloses limitation [1.5]. Decl. ¶¶ 174-176.

Grotmol discloses that its processor or a feature of an application running thereon receives additional digital pictures from the camera over time. Ex. 1004, 38:36-39 (“In some implementations, the robot may comprise a robotic vehicle (e.g., 160 in FIG.1) comprising a camera (e.g., 166). The robot may comprise a controller configured to store a *sequence of images* obtained during training phase.”); 39:23-

28 (“During autonomous operation, control process of the robot may be configured to compare a given (e.g., the most recent, current) image with *one or more* the images from the training buffer.”). In Grotmol, “the most recent, current” image corresponds to the claimed “second circumstance representation” and it is compared to the stored images in the training buffer (the claimed first circumstance representation). Decl. ¶ 175. Furthermore, Grotmol discloses that the “most recent, current” image (*i.e.*, second circumstance representation) can be detected by a camera of the robot that was used during the training mode (*i.e.*, first device), or it can be detected by the camera of a different robot (*i.e.*, second device). Decl. ¶¶ 174-175; *see also* Ex. 1004, 33:27-32 (“The trained configuration may be stored. A saved configuration may be loaded into *the robot being trained* thereby providing one or more trained behaviors to the robot. In some implementations, the trained configuration may be loaded to *one or more other robots* in order to provide learned behaviors.”).

Accordingly, Grotmol discloses generating or receiving a second circumstance representation detected by a sensor of the first device or a second device. Decl. ¶¶ 174-176.

- f. **[1.6] anticipating the first one or more instruction sets for operating the first device based on at least partial match between the second circumstance representation and the first circumstance representation;**

Grotmol discloses limitation [1.6]. Decl. ¶¶ 177-178.

Grotmol discloses that “motor commands at a given time step may be obtained by taking the stored motor commands from the training buffer that may correspond to the *best matching* stored image.” Ex. 1004, 43:5-9. Grotmol discloses that the matching step may comprise calculating a similarity measure between a current image and a stored image, then determining whether that similarity measure breaches a threshold (*i.e.*, “at least partial match between the second circumstance representation and the first circumstance representation”). Ex. 1001, 169:64-67.

The method may comprise, based on the comparison, determining a *similarity measure* for a given training set of the plurality of training feature sets. The similarity measure may characterize similarity between features of the subset and features of the given training set. The method may comprise, responsive to the similarity measure *breaching a threshold*, selecting one or more training sets from the plurality of training sets. The method may comprise determining one or more potential control signals for the robot.

Ex. 1004, 2:29-38; *see also id.*, 39:13-41, 40:64-41:18. As used in Grotmol, “the term feature may be used to describe one or more integer or floating point values characterizing the input, e.g., the presence or absence of an edge, corner, shape, texture, color, object, at particular locations in the image, *values of pixels in an image*, patches of color texture, brightness in the image, and/or in the image as a whole” Ex. 1004, 28:19-24. As used in Grotmol, a “training set” may comprise “an instance of the video frame, an instance of the binary signal, and an instance of

a motor control indication configured to cause the robot to execute an action.” Ex. 1004, 3:35-39.

Both Grotmol and the ’344 patent disclose that determining a “partial match” may comprise pixel-wise difference analysis between two digital pictures. *Compare* Ex. 1001, 66:34-38 (discussing “matching”), 73:24-56 (discussing object detection and recognition by “comparing regions of pixels”) *with* Ex. 1004, 3:13-15 (“The first input type may comprise a digital image comprising a plurality of pixel values.”), 41:8-17 (“A similarity metric . . . may be determined by performing pixel-wise difference between the overlapping regions (e.g., the frames 2020, 2040 in FIG. 20B) of the given image and the matching image.”).

Accordingly, Grotmol discloses anticipating an instruction set for operating the first device learned in the learning process based on a partial match between the second circumstance representation and the first circumstance representation. Decl. ¶¶ 177-178.

- g. **[1.7] and at least in response to the anticipating, executing the first one or more instruction sets for operating the first device, wherein the first device or the second device autonomously performs one or more operations defined by the first one or more instruction sets for operating the first device.**

Grotmol discloses limitation [1.7]. Decl. ¶¶ 179-182.

As described above for limitation [1.6], Grotmol discloses anticipating the learned instruction sets for operating the device by comparing a new image with a

stored image at least by matching portions of the images and determining whether those matching portions exceed a threshold number. Grotmol further discloses “executing the first one or more instruction sets for operating the first device learned in the learning process” in response to the “anticipating.” Grotmol describes that after it is determined that the similarity measure between the new image and image from the training set breaches a threshold, one or more potential control signals corresponding to the image from the training set that were obtained during a training operation of the robot (the instruction sets correlated with the first digital picture) are determined and executed.

The method may comprise, responsive to the similarity measure breaching a threshold, selecting one or more training sets from the plurality of training sets. The method may comprise determining one or more potential control signals for the robot. Individual ones of the one or more potential control signals may be associated with a corresponding training set of the plurality of training sets. . . . In some implementations, the control signal may be configured *to cause the robot to execute the action*.

Ex. 1004, 2:34-41, 3:12-13; *see also id.*, 43:4-11 (“In some implementations, motor commands at a given time step may be obtained by taking the stored motor commands from the training buffer that may correspond to the best matching stored image.”).

Grotmol discloses that the robot (first device) autonomously performs the operations defined by the first instruction set for operating the device. Ex. 1004, 34:56-62 (“In some implementations, the trainer may be provided with the control

of the robot during training. Upon completion of the training, the robot may be configured to operate autonomously.”); Abstract (“During autonomous operation, an acquired image may be compared with one or more images from the training buffer in order to determine the most likely match.”); 43:4-11 (“In some implementations, motor commands at a given time step may be obtained by taking the stored motor commands from the training buffer that may correspond to the best matching stored image.”). Grotmol further discloses that training sets for performing autonomous operation may be loaded into the robot used for training (*i.e.*, first device), or may be loaded into another robot (*i.e.*, second device), and therefore, discloses that a first device or a second device autonomously performs the operations defined by the instructions stored in the training set. Ex. 1004, 33:21-32.

Accordingly, Grotmol discloses executing the instruction set for operating the first device learned in the learning process, where the first device or the second device autonomously performs an operation defined by the instruction set. Decl. ¶¶ 179-182.

2. Claim 2

Grotmol discloses claim 2 that depends from claim 1 and requires that “wherein the first one or more instruction sets for operating the first device include one or more information about one or more states of: the first device, or a portion of the first device.” Decl. ¶ 183.

Grotmol discloses training sets (instruction sets) that contain state information of the first device. As described above for limitation [1.3], Grotmol discloses accessing a memory component configured to store training sets comprising “an instance of the video frame, an instance of the binary signal, and an instance of a motor control indication configured to cause the robot to execute an action.” Ex. 1004, 3:36-39. Grotmol further discloses that “[a]dditional sensory information (e.g., vehicle motion information, ambient environment information, vehicle operational parameters) corresponding to time instances when the images are acquired may be stored,” which constitutes information about the state of the robot (*i.e.*, first device). *Id.*, 39:17-22.

3. Claim 3

Grotmol discloses claim 3 that depends from claim 1 and requires “wherein the first circumstance representation includes a first one or more object representations, and wherein the second circumstance representation includes a second one or more object representations.” Decl. ¶¶ 184-188. As described above for limitations [1.3] and [1.4], Grotmol discloses the claimed first and second circumstance representations. Decl. ¶¶ 165-173.

The '344 patent does not define “object representations” beyond its use in the phrase “collection of object representations,” which it defines as “one or more representations of objects detected by the sensor at a time.” Ex. 1001, 3:42-44. The

“sensor” that detects the object representations may include a “[c]amera ... for capturing one or more pictures” or a “motion picture camera that can capture streams of pictures (i.e. motion pictures, videos, etc.).” *Id.*, 69:52-63. The ’344 patent provides examples of objects represented in pictures such as “person, cat, vehicle, building, street, tree, rock, etc.” Ex. 1001, 72:59-64. The ’344 patent further describes various approaches to object recognition, such as “identifying and/or analyzing features such as lines, edges, ridges, corners, blobs, regions, and/or their relative positions, sizes, shapes, etc.” *Id.*, 75:8-12.

Similarly, Grotmol discloses that images 2000 and 2010 may “comprise representations of visual field that may be obtained at times $t_1 < t_2$, respectively.” Ex. 1004, 38:54-56. These same images 2000 and 2010 may “comprise *representations [of] one or more objects*, e.g., a tree 2002, a rectangle, a ball 2008, and/or other objects.” Ex. 1004, 38:56-65; *see also id.*, 38:65-39:12 (referring to objects shown within images 2000 and 2010 as “object representations” and “representations”).

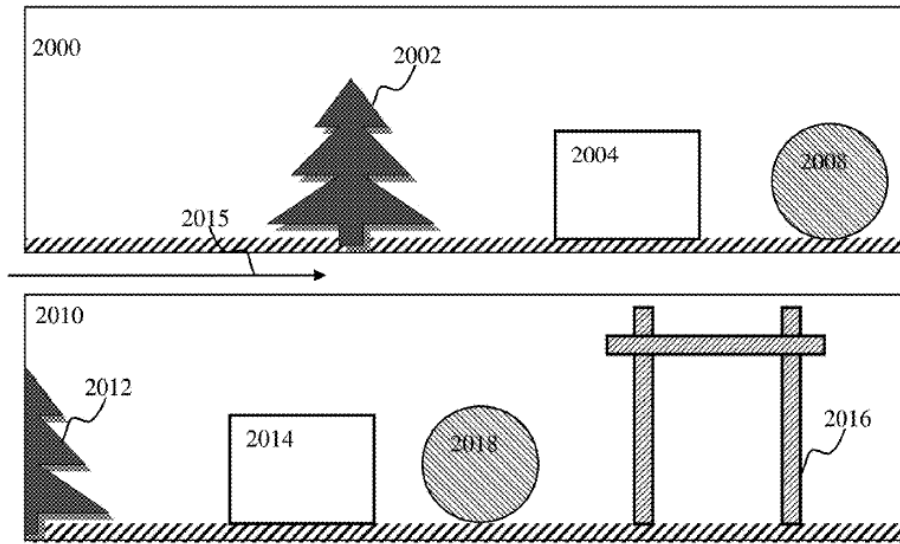


FIG. 20A
Ex. 1004, Fig. 20A

Grotmol likewise discloses that the sensor used to capture these images “may provide information associated with objects within the field-of-view In one or more implementations, such as *object recognition*, and/or obstacle avoidance, the output provided by the sensor component 166 may comprise a stream of pixel values associated with one or more digital images.” Ex. 1004, 8:57-9:13. *See, e.g., id*, 9:41-52 (“In one or more implementations, object recognition and/or classification may be implemented”). Accordingly, Grotmol discloses the first and second circumstance representations including one or more object representations. Decl. ¶ 184-188.

4. Claim 4

Grotmol discloses claim 4 that depends from claim 3 and requires “wherein the second circumstance representation represents the second circumstance detected

at least in part by the one or more sensors of the first device, and wherein the first device autonomously performs the one or more operations defined by the first one or more instruction sets for operating the first device.” Decl. ¶ 189. As described above with respect to limitations [1.5] and [1.7] of claim 1, Grotmol discloses a second circumstance representation is detected by a camera of the robot used for training (*i.e.*, first device), and the same robot autonomously performs the operations defined by the instructions stored in the training set. Decl.¶ 173-176, 179-182.

5. Claim 5

Grotmol discloses claim 5 that depends from claim 3 and requires “wherein the second circumstance representation represents the second circumstance detected at least in part by the one or more sensors of the second device, and wherein the second device autonomously performs the one or more operations defined by the first one or more instruction sets for operating the first device.” Decl. ¶ 190. As described above with respect to limitations [1.5] and [1.7] of claim 1, Grotmol discloses that another robot (*i.e.*, second device) can be loaded with a training set created by operating a first robot (*i.e.*, first device) in order to perform autonomous operation of the another robot (*i.e.*, second device). Decl. ¶¶ 174-176, 179-182. Thus, Grotmol discloses a second circumstance representation is detected by a camera of another robot (*i.e.*, second device) wherein the another robot (*i.e.*, second device)

autonomously performs operations defined by instructions for operating the first robot (*i.e.*, first device). Decl. ¶ 190.

6. Claim 7

- a. **[7.1] The system of claim 3, wherein the one or more processors are further configured to perform at least: modifying: the first one or more instruction sets for operating the first device, or a copy of the first one or more instruction sets for operating the first device,**

Grotmol discloses limitation [7.1]. Decl. ¶¶ 191-193.

Grotmol discloses insertion of user specified instructions into the training sets that modify existing instruction sets. Ex. 1004, 34:49-55 (“In some implementations of supervised training of robots, control instructions (also referred to as corrections) produced by the trainer (e.g., human) may be combined with control instructions produced by the robot controller instructions (predictions).”). Additionally, Grotmol discloses that autonomous operation of a robot may include modifying the instruction set to compensate for example, “position mismatch” and “orientation mismatch.” *Id.*, 42:46-43:3; *see also* Abstract (“An evaluation may be performed in order to determine if the image may correspond to a shifted (e.g., left/right) version of a stored image as previously observed. If the new image is shifted left, right turn command may be issued. If the new image is shifted right then left turn command may be issued.”). For example, Grotmol discloses that the “motor commands at a given time step may be obtained by taking the stored motor commands from the

training buffer that may correspond to the best matching stored image,” and “[t]hose motor commands may be combined with the output from the PID controller in order to stabilize operation of the robot.” *Id.*, 43:5-11. Thus, Grotmol discloses that the “first one or more instruction sets for operating the first device,” are “modified.” Decl. ¶¶ 191-193.

- b. [7.2] and wherein the anticipating the first one or more instruction sets for operating the first device based on the at least partial match between the second circumstance representation and the first circumstance representation includes: anticipating the modified the first one or more instruction sets for operating the first device based on the at least partial match between the second circumstance representation and the first circumstance representation, or anticipating the modified the copy of the first one or more instruction sets for operating the first device based on the at least partial match between the second circumstance representation and the first circumstance representation, and

Grotmol discloses limitation [7.2]. Decl. ¶ 194.

As described above for limitation [1.6], Grotmol discloses that an instruction set is selected based on a partial match between a stored image (first circumstance representation) and the current, most recent image (second circumstance representation). Decl. ¶¶ 177-178; Ex. 1004, 2:29-38, 28:19-24, 34:49-55, 39:13-41, 40:64-41:18, 42:46-43:3, 43:5-11. And as described above for limitation [7.1], Grotmol discloses modifying an instruction set for operating the first device. A POSITA would understand that where a partial match occurs with a stored image

that is associated with an instruction, and the instruction is modified, the anticipated instruction would be the modified instruction. Decl. ¶¶ 191-193. Accordingly, Grotmol discloses anticipating a modified instruction set based on at least a partial match between the second circumstance representation and the first circumstance representation. Decl. ¶ 194.

- c. **[7.3] wherein the executing the first one or more instruction sets for operating the first device includes: executing the modified the first one or more instruction sets for operating the first device, or executing the modified the copy of the first one or more instruction sets for operating the first device, and wherein the autonomously performing, by the first device or by the second device, the one or more operations defined by the first one or more instruction sets for operating the first device includes: autonomously performing, by the first device or by the second device, one or more operations defined by the modified the first one or more instruction sets for operating the first device, or autonomously performing, by the first device or by the second device, one or more operations defined by the modified the copy of the first one or more instruction sets for operating the first device.**

Grotmol discloses limitation [7.3]. Decl. ¶ 195.

As described above for limitations [1.6], [7.1], and [7.2], Grotmol discloses that an instruction set is selected based on a partial match between a stored image (first circumstance representation) and the current, most recent image (second circumstance representation), and that where a partial match occurs with a stored image that is associated with an instruction, and the instruction has been modified,

the anticipated instruction would be the modified instruction. Decl. ¶¶ 177-178, 191-194. Furthermore, as described above for limitation [1.7], Grotmol discloses executing the anticipated instruction set, and the first device or the second device autonomously performs operations defined by the instruction set. Decl. ¶¶ 179-182. A POSITA would understand that where the anticipated instruction is modified, the modified instruction would be executed and the first device or second device would perform the operations defined by the modified instruction set. Decl. ¶ 195.

7. Claim 8:

- a. **[8.1] The system of claim 3, wherein the one or more processors are further configured to perform at least: modifying at least one of: the first circumstance representation, a copy of the first circumstance representation, the second circumstance representation, or a copy of the second circumstance representation,**

Grotmol discloses limitation [8.1]. Decl. ¶ 196.

Grotmol discloses that pixels from individual image instances are selected, scaled, normalized, and/or equalized. *See* Ex. 1004, 4:18-22 (“The method may comprise applying a transformation to an instance of the sensory input. The transformation may be configured to produce scaled input based on analysis of the history.”); 30:47-31:2 (discussing “normalizing” sensory input by “shifting and/or scaling input features”); 31:61-32:5 (discussing equalizing sensory data); 3:42-46;

4:58-63; 31:3-8. Thus, Grotmol discloses a system with processors configured to modifying the first and second circumstance representations. Decl. ¶ 196.

- b. **[8.2] and wherein the at least partial match between the second circumstance representation and the first circumstance representation includes: (i) at least partial match between the modified the second circumstance representation and the first circumstance representation, ... (iii) at least partial match between the second circumstance representation and the modified the first circumstance representation, ... [or] (v) at least partial match between the modified the second circumstance representation and the modified the first circumstance representation, ...**

Grotmol discloses limitation [8.2]. Decl. ¶ 197.

As described with respect to limitations [1.6] and [8.1], Grotmol discloses anticipating an instruction set based on at least a partial match between a current image and a stored image of a training set, and also that captured images may be modified. Decl. ¶¶ 177-178, 196. A POSITA would understand that where either the current image or the stored image have been modified, the partial match would be determined using the modified current image (*i.e.*, modified second circumstance representation) or the modified stored image (*i.e.*, modified first circumstance representation). Decl. ¶ 197. Thus, Grotmol discloses a partial match between: (i) a modified second circumstance representation and a first circumstance representation, (ii) a second circumstance representation and a modified first

circumstance representation, and (iii) a modified second circumstance representation and a modified first circumstance representation.

8. Claim 9

Grotmol discloses claim 9 that depends from claim 3 and requires “wherein the knowledgebase further includes a third circumstance representation correlated with a second one or more instruction sets for operating the first device, and wherein the third circumstance representation represents a third circumstance detected at least in part by the one or more sensors of the first device, and wherein at least a portion of the third circumstance representation or at least a portion of the second one or more instruction sets for operating the first device is learned in another learning process that includes operating the first device at least partially by the user.” Ex. 1001, claim 9; Decl. ¶ 198.

As described above with respect to claim 1, Grotmol discloses that multiple images may be acquired, with images respectively stored in association with corresponding motor commands issued at the time the image was acquired. Grotmol further discloses that multiple training processes can be conducted to learn multiple tasks. Ex. 1004, 24:17-67 (describing training of “simple tasks” and “complex tasks,” including behaviors such as “fetch” and “grasping.”); 26:1-20 (describing that “[t]he system may be configured to learn the *second* behavior from the user.”). Thus, Grotmol discloses the knowledgebase includes a first circumstance

representation (as explained for limitation [1.3]), and additional circumstance representations (*i.e.*, claimed third circumstance representation), where the additional circumstance representation is detected by sensor of the first device and is learned in *another* learning process that includes operating the device by the user. Decl. ¶¶ 165-172; 198.

9. Claim 13

Grotmol discloses Claim 13 that depends from claim 3 and requires the “learning” of the first circumstance representation or the first instruction set, and the “anticipating” of the first instruction set is done at least in part by an “artificial intelligence system.” Decl. ¶¶ 199-202.

The '344 patent broadly describes an “artificial intelligence unit” as a generic piece of hardware or software. *See e.g.*, Ex. 1001, 3:1-41. Neither claim 13 nor the specification of the '344 patent limits the artificial intelligence system to a specific artificial intelligence model or technique. Ex. 1001, 103:33-48. The '344 patent identifies “supervised learning” and “unsupervis[ing] learning” as examples of artificial intelligence models or techniques that can be used, and for anticipating based on a partial match “any level of similarity, however high or low, maybe be utilized” and “any one or more properties (*i.e.*, existence, type, identity, distance, bearing/angle, location, shape/size, activity, etc.) of one or more objects can be

utilized for determining similarity or match of circumstances.” Ex. 1001, 103:33-48, 110:28-44, 111:10-15.

As described above for limitation [1.3], Grotmol discloses a system in which processors perform the operation of learning correlations between circumstance representations (images captured by the robotic device) and instruction sets for operating the first device (motor commands). Decl. ¶¶ 165-172. Grotmol further describes that either supervised or unsupervised learning may be utilized. Ex. 1004, 16:8-12 (“the control entity may comprise a computerized agent such as a multifunction adaptive controller operable using reinforcement and/or *unsupervised learning* and capable of training other robotic devices for one and/or multiple tasks.”); 24:17-20 (“Apparatus and methods for hierarchical learning are disclosed herein, in accordance with one or more implementations. In *supervised learning*, a user may train simple tasks by demonstrating example tasks to the system.”). And as described above for limitation [1.6], Grotmol discloses a system in which one or more processors perform the operation of anticipating an instruction by calculating a similarity measure between a current image and a stored image, then determining whether that similarity measure breaches a threshold. Decl. ¶¶ 177-178. Accordingly, Grotmol discloses an artificial intelligence system that performs the learning and anticipating steps. Decl. ¶¶ 199-202.

10. Claim 14

- a. **[14.1] The system of claim 13, wherein the artificial intelligence system includes: one or more inputs for receiving one or more object representations, and one or more outputs for providing one or more instruction sets,**

Grotmol discloses limitation [14.1]. Decl. ¶¶ 203-204.

As described above for limitations [1.3], [1.6], claim 3, and claim 13, Grotmol discloses an artificial intelligence system utilizing one or more processors that perform the operation of learning to correlate images including object representations with instructions, and anticipating an instruction for operating a robot device based on a partial match between a current image and a stored image. Decl. ¶¶ 165-172, 177-178, 184-188, 199-202. As shown in Figure 8 of Grotmol, the processor is connected to the “sensor” functionality to receive as input images including one or more object representations. *See* Ex. 1004, 9:1-3 (“[T]he output provided by the sensor component 166 may comprise a stream of pixel values associated with one or more digital images.”), 38:52-58.

Grotmol further describes that “motor commands at a given time step may be obtained by taking the stored motor commands from the training buffer that may correspond to the best matching stored image” and “[t]hose motor commands may be combined with the output from the PID controller in order to stabilize operation of the robot.” *Id.*, 43:4-11. For example, as shown in Figure 9, there is an output for

“control signal 420 [that] may comprise one or more motor commands (e.g., pan camera to the right, turn right wheel forward), sensor acquisition commands (e.g., use high resolution camera mode), and/or other commands.” *Id.*, 15:26-33. Thus, Grotmol discloses inputs for receiving object representations and outputs for providing instruction sets. Decl. ¶¶ 203-204.

- b. **[14.2] and wherein the learning, at least in part by the artificial intelligence system, the at least the portion of the first circumstance representation or the at least the portion of the first one or more instruction sets for operating the first device includes applying the first one or more object representations to the one or more inputs of the artificial intelligence system,**

Grotmol discloses limitation [14.2]. Decl. ¶ 205.

As explained above with respect to claim 3 and limitation [14.1], Grotmol discloses that the learning process includes recognizing objects within a given image captured by a camera, and further that the processor includes an input for receiving images having one or more object representations. Decl. ¶¶ 184-188. Thus, Grotmol discloses that a learning process includes applying one or more object representations to the input of an artificial intelligence system. Decl. ¶ 205.

- c. [14.3] and wherein the anticipating, at least in part by the artificial intelligence system, the first one or more instruction sets for operating the first device includes: applying the second one or more object representations to the one or more inputs of the artificial intelligence system; and receiving the first one or more instruction sets for operating the first device from the one or more outputs of the artificial intelligence system.

Grotmol discloses limitation [14.3]. Decl. ¶ 206.

As described above with respect to limitations [1.3] and [1.6], Grotmol discloses a processor that performs the step of anticipating an instruction for operating a first device by comparing a new image (second one or more object representations) with a stored picture at least by matching portions of the pictures and determining whether those matching portions exceed a threshold number to output a control signal. Decl. ¶¶ 165-172, 177-178. Thus, Grotmol discloses applying a new image (second one or more object representations) to an input of an artificial intelligence system running on the processor, and receiving from an output of an artificial intelligence system running on the processor a motor command (instruction set) for operating the first device. Decl. ¶ 206.

11. Claim 15

Grotmol discloses claim 15 that depends from claim 3 and requires “the artificial intelligence system includes at least one of: the knowledgebase, or a neural network.” Decl. ¶ 207.

The '344 patent describes that an artificial intelligence unit may be coupled to a memory unit and the knowledgebase may be stored in the memory unit. Ex. 1001, 3:1-3, 16:39-40. As described above for claim 13, Grotmol discloses an artificial intelligence system running on a processor that performs the learning and anticipating steps, and further as described limitations [1.2], [1.3] and [1.4], Grotmol discloses the claimed knowledgebase stored in a memory unit that is coupled to the processor. Thus, Grotmol discloses an artificial intelligence system that includes a knowledgebase. Decl. ¶¶ 164-173, 199-202, 207.

12. Claim 16

Grotmol discloses claim 16 that depends from claim 3 and requires “wherein the first circumstance representation correlated with the first one or more instruction sets for operating the first device include at least a portion of the first one or more object representations connected, using at least one or more connections, with at least a portion of the first one or more instruction sets for operating the first device.” Decl. ¶ 208.

As described above with respect to claim 3, Grotmol discloses storing connected object representations and instruction sets as trained behaviors. Ex. 1004, 33:24-27 (“The trained controller of the robot may comprise a trained configuration configured to enable autonomous operation (e.g., without user input) by the robotic device.”). Decl. ¶¶ 184-188. The trained behaviors “may be configured based on

learning of associations between sensory context (e.g., presence of an obstacle in front of the robotic vehicle) and a respective action (e.g., right turn) during training.” *Id.*, 33:59-62. An association between an object representation and an action is a “connection” between the two. Decl. ¶ 208. Thus, Grotmol discloses a system where a first circumstance representation correlated with a first instruction includes a connection between an object representation and an instruction. *Id.*

13. Claim 17

Grotmol discloses claim 17 that depends from claim 3 and requires “wherein the anticipating the first one or more instruction sets for operating the first device based on the at least partial match between the second circumstance representation and the first circumstance representation includes determining the first one or more instruction sets for operating the first device based on at least partial match between the second one or more object representations and the first one or more object representations.” Decl. ¶ 209.

As described above with respect to limitation [1.6] and claim 3, Grotmol discloses anticipating an instruction set based on a partial match between the first and second circumstance representation, and further that the first and second circumstance representations include object representations. Decl. ¶¶ 177-178, 184-188. Grotmol further discloses that determining a partial match includes determining a partial match between object representations. *See e.g.*, Ex. 1004, 38:52-39:28

(describing how representations of objects within two images may be compared);

Decl. ¶ 209.

14. Claim 18

- a. **[18.1] The system of claim 3, wherein at least some elements of the system are included in: a single device, or multiple devices, and**

Grotmol discloses using a “self-contained robotic apparatus,” and thus, discloses that at least some of the elements of the system are included in a single device. Ex. 1004, 3:27-30; *see supra* limitation [1.1]; Decl. ¶¶ 163, 210.

- a. **[18.2] wherein the one or more processors include: one or more microcontrollers, one or more computing devices, or one or more electronic devices, and**

Grotmol discloses one or more processor circuits include a microcontroller, computing circuits or electronic circuits. *See supra* limitation [1.2]; Decl. ¶¶ 164, 211.

- b. **[18.3] wherein the memory includes: a non-transitory machine readable medium, a volatile memory, a non-volatile memory, a storage device, or a storage system, and**

Grotmol discloses that its “memory component” may comprise a “non-volatile memory.” Ex. 1004, 38:39-42 (“In some implementations, the images may be stored in an ordered buffer in a non-volatile memory.”); *see supra* limitation [1.3]; Decl. ¶¶ 165-172, 212.

- c. **[18.4] wherein the one or more sensors of the first device include at least one of: one or more cameras... and wherein the one or more sensors of the second device include at least one of: one or more cameras, ...and**

As described above for limitation [1.3] and claim 3, Grotmol discloses that a robotic vehicle includes a sensor comprising a camera for detecting objects or object properties. *See e.g.*, Ex. 1004, 8:57-9:3, 10:15-26, 38:33-58, 39:13-22. Furthermore, as described above for claim 5, Grotmol discloses a second robotic vehicle that includes a camera for detecting objects or object properties. *See e.g., id.*, 33:21-35. Decl. ¶¶ 165-172, 184-188, 190, 213-214.

- d. **[18.5] wherein the first device includes: a robot, a vehicle, ... or a mechanical machine, and**

Grotmol discloses that a first device that may be used for training and autonomous operation is a robot (“self-contained robotic apparatus”) that may include a vehicle. Ex. 1004, 3:27-30; *id.*, 38:36-39 (“In some implementations, the robot may comprise a robotic vehicle (e.g., 160 in FIG.1). . . .”); *see supra* claim 1; Decl. ¶¶ 163-182, 215.

- e. **[18.6] wherein the second device includes: a robot, a vehicle, ... or a mechanical machine, and**

Grotmol discloses limitation [18.6]. Decl. ¶¶ 216.

Grotmol discloses that the training sets and training configurations from a first robot (which include the instruction sets for operating first device) can be “stored” or “saved.” Ex. 1004, 33:12-58 (discussing saved training configurations). The saved

training configuration may then be “loaded into the robot being trained thereby providing one or more trained behaviors to the robot.” *Id.*, 33:27-30. Further, the “trained configuration may be loaded to one or more *other robots* in order to provide learned behaviors.” *Id.*, 33:30-32. Thus, Grotmol discloses a second device (“one or more other robots”) that can autonomously perform one or more operations defined by one or more instruction sets for operating the first device (training configuration from first robot (e.g., training sets with circumstance representations correlated to instruction sets) that is loaded into a second one or more other robots to “provide learned behaviors.” Decl. ¶ 216.

f. **[18.7] wherein the first circumstance representation is a data structure that includes one or more data about the first circumstance, and wherein the second circumstance representation is a data structure that includes one or more data about the second circumstance,**

Grotmol discloses limitation [18.7]. Decl. ¶¶ 217-220.

The '344 patent discusses that “any variety of data structures can be used such as, for example, array, list, linked list, doubly linked list, queue, tree, heap, graph, map, grid, matrix, multi-dimensional matrix, table, database, database management system (DBMS), file, neural network, and/or any other type or form of a data structure including a custom one.” Ex. 1001, 65:55-63. As discussed above with respect to limitation [1.3] of claim 1, Grotmol discloses circumstance representations comprising video frame instances. Grotmol discloses that a video frame

“compris[es] a plurality of pixels.” Ex. 1004, 3:31-33, Decl. ¶¶ 165-172. Thus, the video frames disclosed by Grotmol comprise “data structures” in the context of the ’344 patent because they comprise at least an array, map, grid, or other form of pixels. Decl. ¶¶ 218-219.

Further, the entire video frame itself also comprises a “data structure” within the context of the ’344 patent because it is file including image data. The ’344 patent explains that a data structure can comprise a single file or field of data. Ex. 1001, 65:61-65 (“A data structure may include *one* or more fields or data fields that are part of or associated with the data structure. A field or data field may include *a data*, an object, a data structure, and/or any other element or a reference/pointer thereto.”). Further, the ’344 patent states that data structures “can be stored in one or more memories, files, or other repositories,” (*id.*, 65:65-67) just as the video frame instances are stored in the “memory component” of Grotmol. Ex. 1004, 3:27-39.

The video frame instances of Grotmol include “one or more data” about the “circumstance[s] of the first device.” Decl. ¶ 220. The video frame instances are detected by a sensor attached to the first device, and capture images of the devices surroundings for use in object detection and avoidance. Ex. 1004, 8:57-9:3, 10:15-26, 38:52-58.

- g. [18.8] and wherein the first circumstance includes one or more objects detected ... by the one or more sensors of the first device ... during a first time period, and wherein the second circumstance includes: one or more objects detected ... by the one or more sensors of the first device ... during a second time period, or one or more objects detected ... by the one or more sensors of the second device ... during a second time period, and

Grotmol discloses limitation [18.8]. Decl. ¶¶ 221-225.

As described above for limitations [1.3] and [1.5], Grotmol discloses a knowledgebase including a first circumstance representation detected by a first device during a training session (i.e., a first time period), and a second circumstance representation detected by a first or second device in the context of autonomous operation (i.e., a second time period). Also described above for claim 3, Grotmol discloses that the first and second circumstance representations include detected objects. Decl. ¶¶ 165-172, 174-177, 184-188.

The '344 patent provides examples of objects in pictures such as “person, cat, vehicle, building, street, tree, rock, etc.” Ex. 1001, 72:59-64. The '344 patent further describes various approaches to object recognition, such as “identifying and/or analyzing features such as lines, edges, ridges, corners, blobs, regions, and/or their relative positions, sizes, shapes, etc.” *Id.*, 75:8-12.

Similarly, Grotmol discloses that images 2000 and 2010 may “comprise representations of visual field that may be obtained at times $t_1 < t_2$, respectively.” Ex.

1004, 38:54-56. These same images 2000 and 2010 may “comprise *representations* [of] one or more objects, e.g., a tree 2002, a rectangle, a ball 2008, and/or other objects.” *Id.*, 38:56-65.

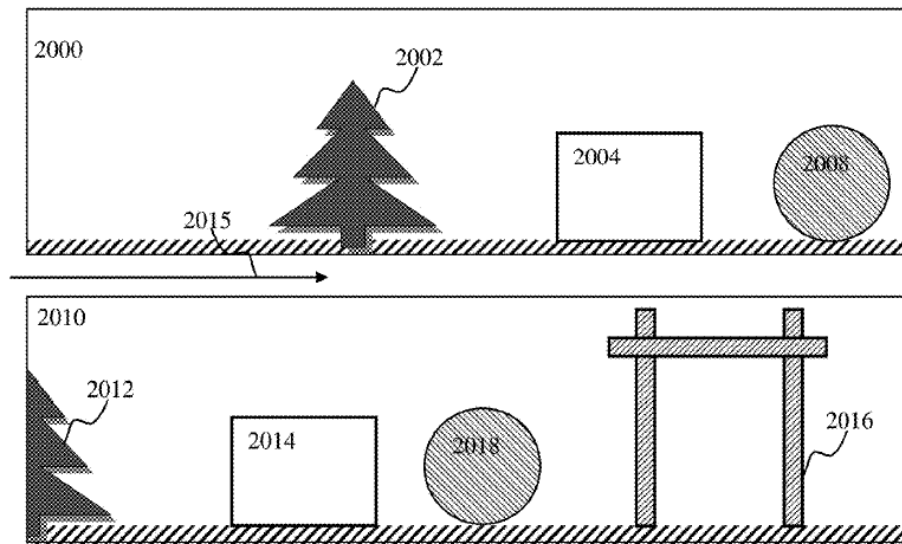


FIG. 20A
Ex. 1004, Fig. 20A

Grotmol likewise discloses that the sensor used to capture these images “may provide information associated with objects within the field-of-view In one or more implementations, such as *object recognition*, and/or obstacle avoidance, the output provided by the sensor component 166 may comprise a stream of pixel values associated with one or more digital images.” Ex. 1004, 8:57-9:13. *See, e.g., id.*, 9:41-52 (“In one or more implementations, object recognition and/or classification may be implemented”). Grotmol, like the ’344 patent, further includes edge detection. *Id.*, 7:60-63 (“[T]he term feature may refer to a representation of an object edge”).

Grotmol describes that the autonomous operation of a robot may be based in part on recognized objects and their characteristics. *Id.*, 10:23-34 (noting that “tasks of the robot may be configured based on a context” such as “state of the environment (e.g., object size, location)”), 24:31-25:5 (describing embodiment in which a robot is trained to “find an object, grasp it, and bring it back to base”).

- h. **[18.9] wherein an instruction set of the first one or more instruction sets for operating the first device includes at least one of: only one instruction, multiple instructions, ... one or more commands, ... or one or more information, and**

As described above for limitation [1.3], Grotmol discloses an instruction set for operating the first device, where the instruction set includes, for example, one or more instructions, one or more commands, and one or more information. *See* Ex. 1004, 3:35-39, 23:3-11, 39:13-22, 43:4-11; Decl. ¶¶ 165-172, 226.

- i. **[18.10] wherein the first one or more object representations include: one object representation, ... and wherein the second one or more object representations include: one object representation, ... and wherein the first one or more object representations include: ... one or more digital pictures that depict one or more objects, ... and wherein the second one or more object representations include: ... one or more digital pictures that depict one or more objects, ... and wherein the at least the portion of the first circumstance representation includes: at least a portion of one object representation ... and**

As described above for limitation [1.3] and claim 3, Grotmol discloses a first

circumstance representation includes an object representation, and a second circumstance representation includes an object representation, and also discloses the first and second circumstance representations include digital pictures that depict one or more objects. Decl. ¶¶ 165-172, 184-188, 227.

- j. **[18.11] wherein the at least the portion of the first one or more instruction sets for operating the first device includes: at least a portion of one instruction set of the first one or more instruction sets for operating the first device ...**

As described above for limitations [1.3] and [1.4], Grotmol discloses motor commands (i.e., at least a portion of the first one or more instruction sets for operating the first device) are learned in a learning process that includes operating the first device at least partially by a user. Accordingly, Grotmol discloses wherein the at least the portion of the first one or more instruction sets for the operating the first device includes at least a portion of one instruction of the first one or more instruction sets for operating the first device. Decl. ¶¶ 165-173, 228.

- k. **[18.12] and wherein: an object of the first circumstance is the same as an object of the second circumstance, multiple objects of the first circumstance are the same as multiple objects of the second circumstance, all objects of the first circumstance are the same as all objects of the second circumstance, or all objects of the first circumstance are different than all objects of the second circumstance.**

As described above with respect to limitations [1.6] and [18.8], Grotmol discloses comparing a first circumstance and second circumstance to determine whether there is at least a partial match, and further discloses that the first and second circumstance include one or more objects, as well as object detection. Decl. ¶¶ 177-178, 221-225. The object(s) in the first circumstance and the second circumstance can be the same or different. *See e.g.*, Ex. 1004, 38:52-39:12, Fig. 20A, Fig. 20B. Grotmol further discloses that “the similarity [between the images] may be determined by performing pixel-wise difference between the overlapping regions (e.g., the frames 2020, 2040 in FIG. 20B) of the given image and the matching image.” *Id.*, 41:10-13. The images may have one object, multiple or all objects that are the same, or have all different objects. Decl. ¶ 228.

15. Claim 19

Independent claim 19 is a method claim that corresponds to system claim 1. The analysis above explaining how Grotmol discloses the limitations of claim 1 applies to the limitations of claim 19 as set forth in the chart below. Decl. ¶ 230.

Claim 19	Corresponding Analysis for Claim 1
[19.1] A method implemented using a computing system that includes one or more processors, the method comprising:	<i>See</i> analysis for limitations [1.1] and [1.2] above.

Claim 19	Corresponding Analysis for Claim 1
[19.2] accessing a memory that stores at least a knowledgebase that includes a first circumstance representation correlated with a first one or more instruction sets for operating a first device, wherein the first circumstance representation represents a first circumstance detected at least in part by one or more sensors of the first device,	<i>See analysis for limitation [1.3] above.</i>
[19.3] and wherein at least a portion of the first circumstance representation or at least a portion of the first one or more instruction sets for operating the first device is learned in a learning process that includes operating the first device at least partially by a user;	<i>See analysis for limitation [1.4] above.</i>
[19.4] generating or receiving a second circumstance representation, wherein the second circumstance representation represents a second circumstance detected at least in part by: the one or more sensors of the first device, or one or more sensors of a second device;	<i>See analysis for limitation [1.5] above.</i>
[19.5] anticipating the first one or more instruction sets for operating the first device based on at least partial match between the second circumstance representation and the first circumstance representation; and	<i>See analysis for limitation [1.6] above.</i>

Claim 19	Corresponding Analysis for Claim 1
[19.6] at least in response to the anticipating, executing the first one or more instruction sets for operating the first device, wherein the first device or the second device autonomously performs one or more operations defined by the first one or more instruction sets for operating the first device.	<i>See</i> analysis for limitation [1.7] above.

16. Claim 20

Independent claim 20 is a system claim that corresponds to claim 1, except that claim 20 requires “means for processing” whereas claim 1 requires “one or more processors.” The analysis above explaining how Grotmol discloses the limitations of claim 1 applies to the limitations of claim 20 as set forth in the chart below. Decl.

¶ 231

Claim 20	Corresponding Analysis for Claim 1
[20.1] A system comprising:	<i>See</i> analysis for limitation [1.1] above.
[20.2] means for processing; and	Grotmol discloses a system comprising means for processing. As explained above, the corresponding structure for this limitation is a processor. <i>See</i> Section VI.A <i>supra</i> . Grotmol discloses this limitation because it describes that its system includes a processor. <i>See, e.g.,</i> Ex. 1004, claim 1 (requiring one or more processors executing instructions stored by a non-transitory computer-readable storage medium). Grotmol also discloses that its system is a self-contained robot apparatus that includes one or more physical processors. <i>Id.</i> , 3:27-30 (“Another aspect of the disclosure relates to a self-

Claim 20	Corresponding Analysis for Claim 1
	contained robot apparatus. The apparatus may comprise a platform, a first sensor, a second sensor, a memory component, and one or more physical processors.”). Decl. ¶ 164.
[20.3] a memory that stores at least a knowledgebase that includes a first circumstance representation correlated with a first one or more instruction sets for operating a first device, wherein the first circumstance representation represents a first circumstance detected at least in part by one or more sensors of the first device,	<i>See</i> analysis for limitations [1.3] above.
[20.4] and wherein at least a portion of the first circumstance representation or at least a portion of the first one or more instruction sets for operating the first device is learned in a learning process that includes operating the first device at least partially by a user, and	<i>See</i> analysis for limitations [1.4] above.

Claim 20	Corresponding Analysis for Claim 1
<p>[20.5] wherein the means for processing is configured to perform at least: generating or receiving a second circumstance representation, wherein the second circumstance representation represents a second circumstance detected at least in part by: the one or more sensors of the first device, or one or more sensors of a second device;</p>	<p><i>See analysis for limitation [1.5] above.</i></p>
<p>[20.6] anticipating the first one or more instruction sets for operating the first device based on at least partial match between the second circumstance representation and the first circumstance representation; and</p>	<p><i>See analysis for limitation [1.6] above.</i></p>
<p>[20.7] at least in response to the anticipating, executing the first one or more instruction sets for operating the first device, wherein the first device or the second device autonomously performs one or more operations defined by the first one or more instruction sets for operating the first device.</p>	<p><i>See analysis for limitation [1.7] above.</i></p>

B. Ground 2

As set forth below, Grotmol in view of the knowledge of a POSITA renders claims 10 and 11 obvious.

1. Claim 10

Grotmol in view of the knowledge of a POSITA renders obvious claim 10, which requires: (i) learning another correlation between a circumstance representation detected by a sensor of a first device and an instruction set for operating the first device; and (ii) the correlation is learned during an additional learning process that involves a operating the first device by another user. Decl. ¶¶ 233.

As described above in Ground 1, Grotmol discloses that a human trainer (“user”) can operate a robot during a training process for learning correlations. Ex. 1004, 23:3-11, 39:13-22. Furthermore, Grotmol discloses that multiple learning processes can be done to train multiple behaviors for the robot, including “simple tasks” and “complex tasks.” *Id.*, 24:17-67, 26:1-20. Grotmol does not limit the number of users that can train the multiple behaviors for the robot, and a POSITA would have understood that different users can train different behaviors, particularly given that Grotmol describes the training of both simple and complex tasks. Decl. ¶ 233. It would have been obvious to a POSITA that different trainers could be used to train behaviors involving different levels of complexity in order to take advantage

of the varying skills and knowledge of trainers. *Id.* Thus, it would be obvious based on Grotmol in view of the knowledge of a POSITA to have a second user perform an additional training resulting in an additional correlation stored in the knowledgebase. *Id.*

2. Claim 11

Grotmol in view of the knowledge of a POSITA renders obvious claim 11, which requires: (i) learning a correlation between a circumstance representation detected by a sensor of another (third) device and an instruction set for operating the third device; and (ii) the correlation is learned during an additional learning process that involves a operating the other (third) device by the same user or another user. Decl. ¶¶ 234-236.

As described above for claim 10, Grotmol discloses that multiple learning processes can be done to train multiple behaviors, and that it would be obvious that the multiple learning processes could be done by different users operating the robot. Grotmol further discloses that a robot can be loaded with a first trained configuration learned during a training process involving operation of that robot by a user, and the robot can also be loaded with another trained configuration learned during a training process involving operation of a different robot by a user. Ex. 1004, 33:24-35; Decl. ¶ 233. For the same reasons discussed with respect to claim 10, it would be obvious to a POSITA that the training process involving operation of the first robot (resulting

in the first trained configuration), and the training process involving operation of the different robot (resulting in another trained configuration) could be done by different users. Decl. ¶¶ 234-236.

C. Ground 3

Grotmol in view of Hickman renders claims 6, 11, and 12 obvious. Hickman discloses a “shared robot knowledge base [that] is updated based on robot experiences so that any particular robot may benefit from the prior experiences of other robots.” Ex. 1005, Abstract. As detailed below, a POSITA would be motivated to combine Grotmol’s autonomous robotic system with Hickman’s shared robot knowledge base.

1. Motivation to Combine Grotmol and Hickman

A person of ordinary skill in the art would have been motivated to combine the autonomous robotic system of Grotmol with the shared robot knowledge base of Hickman. Decl. ¶¶ 143-150.

Grotmol teaches that the training sets and training configurations from a first robot can be “stored” or “saved.” Ex. 1004, 33:12-58 (discussing saved training configurations). The saved training configuration may then be “loaded into the robot being trained thereby providing one or more trained behaviors to the robot.” *Id.*, 33:27-30. Further, the “trained configuration may be loaded to one or more *other robots* in order to provide learned behaviors.” *Id.*, 33:30-32.

Hickman teaches a shared robot knowledge base that could advantageously be paired with Grotmol's autonomous robotic system in order to further facilitate the sharing of learned behaviors.

The advantages of a shared robot knowledge base 306 multiply as robots are deployed around the world. If information about a particular object is "public" and therefore accessible by all robots, then information that a single robot learns about that object and uploads to the shared robot knowledge base 306 may be accessible by other robots all around the world. In operation, many millions or even many billions of robots may ultimately access the shared robot knowledge base 306. As a result, the shared robot knowledge base 306 enables robots to benefit from the collective information learned by many other robots.

Ex. 1005, 11:31-41. Hickman additionally teaches that "dividing application execution and storage between client devices and the cloud computing system allows more processing to be performed by the cloud computing system, thereby taking advantage of the cloud computing system's processing power and capability, for example." *Id.*, 5:41-46. Hickman further teaches that cloud computing may "free users from certain hardware and software installation and maintenance tasks." *Id.*, 1:38-41.

A POSITA would have been familiar with the teachings of both Grotmol and Hickman and would have readily appreciated the advantages of their combination. Decl. ¶¶ 143-150. Indeed, a 2014 Science Magazine article—which included an interview with the CEO of Brain Corporation, i.e., the assignee of Grotmol—

expressly contemplated the benefits of the kinds of shared robot learning described in Hickman:

Efficient chips can run complex—and normally power-hungry—algorithms that enable robots to learn. The Internet, in turn, will allow far-flung robots to share those lessons and skills. A robot that learns how to pick a strawberry without crushing it, for instance, could uplink that skill for the benefit of its kind around the globe. That means neuromorphic computing could offer robots something far more profound than enhanced perceptual skills. When humans collectively pass along life lessons, we call that culture.

Ex. 1006 at 183; *see also* Decl. ¶ 149. In addition, the two references are directed to the same field of the training and autonomous operation of robots (Decl. ¶¶ 143-148), and Hickman is listed on the face of Grotmol as a cited U.S. patent document. Grotmol and Hickman also describe a similar problem in which a robot learns how to “fetch” from a human. *Compare* Ex. 1005, 23:49-24:25 *with* Ex. 1004, 24:31-48; 27:28-40, Fig. 9. Grotmol describes the task as “quite complex.” Ex. 1004, 24:31-33. Hickman teaches that the fetch task can be “optimiz[ed] . . . based on the collective experience of many different robots.” Ex. 1005, 24:18-25.

The resulting combination of Grotmol’s autonomous robotic system and Hickman’s shared robot knowledgebase would have been expected and predictable. Decl. ¶ 150. Grotmol expressly contemplates the sharing of trained configurations among robots, and Hickman’s expansion upon this idea does not interfere with Grotmol’s autonomous robotic system. Decl. ¶¶ 143-150.

For at least these reasons, it would have been obvious to a POSITA to combine Grotmol and Hickman.

2. Claim 6

Dependent claim 6 adds: (i) a server that receives one of a first circumstance representation or first instruction set from the first device; (ii) the second device receives the first circumstance representation or first instruction set from the server; and (iii) the second device autonomously performs operation defined by the first instruction set. Grotmol in combination with Hickman renders claim 6 obvious. Decl. ¶¶ 239-244.

Grotmol discloses that the training sets and training configurations from a robot (which include the instruction sets for operating first device) can be “stored” or “saved.” Ex. 1004, 33:12-58 (discussing saved training configurations). The saved training configuration may then be “loaded into the robot being trained thereby providing one or more trained behaviors to the robot.” *Id.*, 33:27-30. Further, the “trained configuration may be loaded to one or more *other robots* in order to provide learned behaviors.” *Id.*, 33:30-32. Grotmol also discloses that training sets are used for autonomous operation of the robots by comparing acquired images with images from the stored training set to determine the most likely match and the corresponding instruction to perform. *See e.g., id.*, 33:24-35, Abstract. Thus, Grotmol discloses storing a circumstance representation or first one or more instruction sets for

operating a first device and a second device receiving the first circumstance representation or the first one or more instruction sets for operating the first device. Decl. ¶¶ 239-240.

Grotmol discloses, for example, situations in which the user may wish to train the robot to “find an object, grasp it, and bring it back” (i.e., fetch). Ex. 1004, 24:31-25:5. Grotmol describes that this “task of playing fetch may be quite complex” and outlines numerous user training steps that would be needed in order to teach this skill to the robot. *Id.* A POSITA would understand that it would be advantageous to share learned skills among robots. Decl. ¶ 240.

Hickman discloses such a system wherein learned robot behaviors are uploaded to a shared knowledge base for use in other robots, such as the system described in Hickman. Decl. ¶ 242-243. Hickman discloses “a shared robot knowledge bases for use with cloud computing systems” that “collects data from a robot about an object the robot has encountered in its environment, and stores the received data in the shared robot knowledge base ... sends instructions for interacting with an object to a robot ... sends instructions to a robot for executing an application based on information stored in the shared robot knowledge base.” Ex. 1005, Abstract. Hickman discloses using a cloud computing system, where “[t]he ‘cloud’ may refer to a service or a group of service accessible over a network (e.g., the Internet) by clients, *server devices*, and cloud computing systems...” *Id.*, 5:20-

23. Moreover, “[t]he cloud computing system may be configured to (i) send and receive data and queries to and from multiple robots, (ii) send data retrieved from the shared robot knowledge base to robots, and (iii) update data in the shared robot knowledge base based on data received from robots.” *Id.*, 5:7-11; *See also* Fig. 6.

As noted in Hickman, the “information in the shared robot knowledge bases is updated based on robot experiences so that any particular robot may benefit from prior experiences of other robots.” *Id.*, Abstract. Hickman describes training a robot how to properly grasp objects such as a cup, uploading and storing this information on a shared knowledge base (memory), then having another robot obtain these instructions from the memory and executing the instructions. *Id.*, 26:21-27:21. Hickman further discloses that when a robot queries the cloud processing engine, “the cloud computing system sends data associated with the object of the second query to the second robot in response to the second query received from the second robot at block 606. The data may include...instructions for interacting with the object of the second query.” *Id.*, 29:33-41; *see also* Decl. ¶¶ 242-243.

In the combination of Grotmol and Hickman, the training configuration learned by operation of a robot described in Grotmol would be uploaded to the server containing the “shared robot knowledge base” described in Hickman, and a second robot would receive the circumstance representations and instructions stored in the knowledge base in order to perform autonomous operation. Decl. ¶ 242-244. Thus,

the combination of Grotmol and Hickman discloses (i) receiving, at a server, a first circumstance representation or first instruction set from a first device; (ii) a second device receiving from a server a first circumstance representation or first instruction set; and (iii) the second device autonomously performing an operation defined by the first instruction set. *Id.*

3. Claim 11

Grotmol in view of Hickman renders obvious claim 11, which requires (i) learning a correlation between a circumstance representation detected by a sensor of another (third) device and an instruction set for operating the third device; and (ii) the correlation is learned during an additional learning process that involves a operating the another (third) device by the same user or another user. Decl. ¶¶ 245-247.

As described above for Grounds 1 and 2, Grotmol discloses (i) training a robot wherein images (circumstance representations) detected by a camera during operation of the robot are correlated with motor commands (instructions); (ii) multiple training (learning) processes can be done to train multiple behaviors; and (iii) a robot can be loaded with a first trained configuration learned during a training process involving operation of that robot by a user, and can also be loaded with another trained configuration from a training process involving operation of a different robot by a user. Ex. 1004, 33:24-35; Decl. ¶¶ 163-235.

Hickman discloses “a shared robot knowledge bases for use with cloud computing systems” that “collects data from a robot about an object the robot has encountered in its environment, and stores the received data in the shared robot knowledge base ... sends instructions for interacting with an object to a robot ... sends instructions to a robot for executing an application based on information stored in the shared robot knowledge base.” Ex. 1005, Abstract. As noted in Hickman, the “information in the shared robot knowledge bases is updated based on robot experiences so that any particular robot may benefit from prior experiences of other robots.” *Id.* Hickman discloses that “robot experiences” may include “‘coaching’ from *humans*” (*Id.*, 23:36-48), like the training described in Grotmol. Decl. ¶ 246. For example, similar to the “fetch” behavior described in Grotmol, Hickman describes a human coach training a robot how to properly grasp objects such as a cup, uploading and storing this information on a shared knowledge base (memory), then having another robot obtain these instructions from the memory and executing the instructions. Ex. 1005, 26:21-27:21.

As explained above, it would have been obvious to combine Hickman’s shared robot knowledge base with Grotmol’s autonomous robot system such that the shared knowledgebase includes the training of multiple robots by multiple users. Decl. ¶ 247. The combination discloses a knowledgebase that includes a circumstance representation correlated with instructions for operating a robot (i.e.,

claimed first circumstance correlated with first instruction set) learned during a training using the the robot, and also a circumstance representation correlated with instructions for operating a different robot (i.e., claimed third circumstance representation correlated with second instruction set) learned during a training using the different robot. *Id.*

4. Claim 12

- a. **[12.1] The system of claim 3, wherein the knowledgebase further includes a third circumstance representation correlated with a second one or more instruction sets for operating a third device, and wherein the third circumstance representation represents a third circumstance detected at least in part by one or more sensors of the third device, and wherein the one or more processors are further configured to perform at least:**

Limitation [12.1] includes the same limitations regarding the knowledgebase including an additional circumstance representation correlated with an instruction set as claim 11. Accordingly, Grotmol in combination with Hickman renders obvious limitation [12.1] for the same reasons as described above for claim 11. Decl.

¶¶ 245-247, 248.

- b. [12.2] generating or receiving a fourth circumstance representation, wherein the fourth circumstance representation represents a fourth circumstance detected at least in part by one or more sensors of a fourth device; anticipating the second one or more instruction sets for operating the third device based on at least partial match between the fourth circumstance representation and the third circumstance representation; and at least in response to the anticipating the second one or more instruction sets for operating the third device, executing the second one or more instruction sets for operating the third device, wherein the fourth device autonomously performs one or more operations defined by the second one or more instruction sets for operating the third device.

Grotmol in combination with Hickman renders obvious limitation [12.2] that requires a fourth device detects a circumstance representation, an instruction set is anticipated based on a partial match between the detected circumstance representation and the third circumstance representation stored in the knowledgebase, and executing the instruction set such that the fourth device autonomously performs the operation defined by the instruction set. Decl. ¶ 249.

As described above for claim 11 and limitation [12.1], the combination of Grotmol and Hickman renders obvious a knowledgebase including the third circumstance representation correlated with the second instruction set. Grotmol and Hickman further disclose that the knowledgebase can be shared to perform autonomous operation of any number of other devices. Ex. 1004, 33:27-32; Ex. 1005, 11:31-41; Decl. ¶¶ 245-247, 248. In Grotmol, the autonomous operation of a

robot is performed by acquiring an image (circumstance representation) during operation of the robot, comparing the acquired image to images in the knowledgebase to determine whether there is a match, and executing an operation defined by an instruction correlated with a matched image. Ex. 1004, Abstract, 39:13-41.

Thus, for example, the combination discloses the knowledgebase includes multiple circumstance representations correlated with instructions each learned during a training process involving different robots (i.e., claimed first and third circumstance representations detected during learning processes involved first and third device), and then sharing that knowledgebase to perform autonomous operation of multiple other robots (i.e., claimed second and fourth devices) Accordingly, the combination of Grotmol and Hickman renders obvious anticipating an instruction for operating a fourth device based on a partial match between a circumstance represented detected by sensors of a fourth device and a third circumstance representation stored in the knowledgebase, and then executing operations defined by an instruction set correlated with the third circumstance representation. Decl. ¶ 249.

X. CONCLUSION

For the foregoing reasons, *inter partes* review should be instituted and claims 1-20 of the '344 patent should be cancelled as unpatentable.

Respectfully submitted,

Dated: June 9, 2023

By: /s/ Anish Desai

Anish Desai

Lead Counsel for Petitioner

Registration No. 73,760

Weil, Gotshal & Manges LLP

767 Fifth Ave.

New York, NY 10153

T: 212-310-8000

F: 212-310-8007

CERTIFICATE OF COMPLIANCE

1. The undersigned certifies that this brief complies with the type volume limitations of 37 CFR § 42.24(a)(1)(i). This brief contains 13,922 words (excluding the table of contents, the table of authorities, mandatory notices under 37 CFR § 42.8, the certificate of service, certificate of compliance, and appendix of exhibits), as calculated by the “Word Count” feature of Microsoft Word 2016, the word processing program used to create it.

2. The undersigned further certifies that this brief complies with the typeface requirements of 37 CFR § 42.6(a)(2)(ii) and typestyle requirements of 37 CFR § 42.6(a)(2)(iii). This brief has been prepared in a proportionally spaced typeface using Microsoft Word 2016 in Times New Roman 14 point font.

Respectfully submitted,

Dated: June 9, 2023

By: /s/ Anish Desai
Anish Desai
Lead Counsel for Petitioner
Registration No. 73,760
Weil, Gotshal & Manges LLP
767 Fifth Ave.
New York, NY 10153
T: 212-310-8000
F: 212-310-8007

CERTIFICATE OF SERVICE

Pursuant to 37 C.F.R. § 42.105 and § 42.6(e), the undersigned certifies that on June 9, 2023, a complete and entire copy of this **PETITION FOR *INTER PARTES* REVIEW OF U.S. PATENT NO. 11,238,344, PETITIONER TESLA, INC.'S POWER OF ATTORNEY, and all supporting exhibits** were sent via FedEx next business day delivery to the following:

Jasmin Cosic
Autonomous Devices LLC
2271 Myla Lane
Melbourne, FL 32935

/s/ Timothy J. Andersen/
Timothy J. Andersen
Case Manager
Weil, Gotshal & Manges LLP
2001 M Street, N.W., Suite 600
Washington, D.C. 20036
202-682-7000
timothy.andersen@weil.com