Hochschule
Bonn-Rhein-Sieg
University of Applied Sciences

b-it Bonn-Aachen International Center for Information Technology

Institute for AI and Autonomous Systems

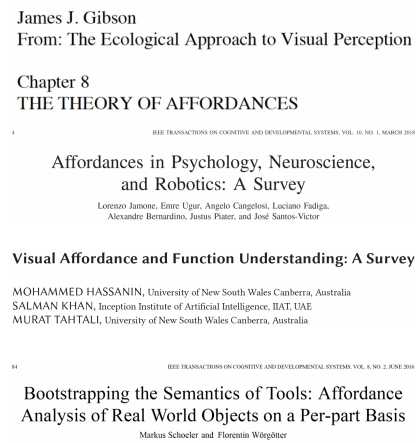
Affordances

An Overview for Cognitive Robotics

Dr. Alex Mitrevski
Master of Autonomous Systems

Structure

- Overview of affordances
- Recognising affordances
- Affordances and robot learning



James J. Gibson
From: The Ecological Approach to Visual Perception

Chapter 8
THE THEORY OF AFFORDANCES

4 IEEE TRANSACTIONS ON COGNITIVE AND DEVELOPMENTAL SYSTEMS, VOL. 10, NO. 1, MARCH 2018

Affordances in Psychology, Neuroscience, and Robotics: A Survey

Lorenzo Jamone, Emre Ugur, Angelo Cangelosi, Luciano Fadiga, Alexandre Bernardino, Justus Piater, and José Santos-Victor

Visual Affordance and Function Understanding: A Survey

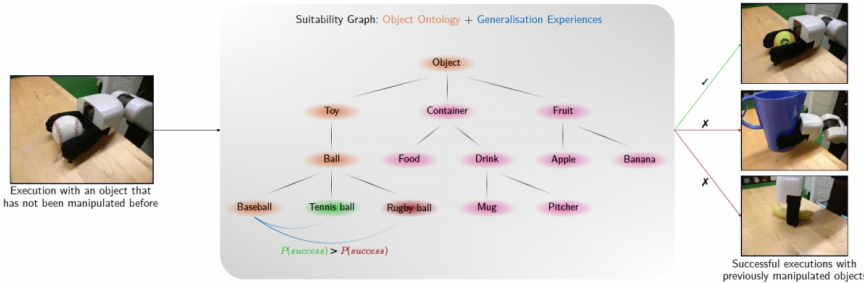
MOHAMMED HASSANIN, University of New South Wales Canberra, Australia
SALMAN KHAN, Inception Institute of Artificial Intelligence, IIAT, UAE
MURAT TAHTALI, University of New South Wales Canberra, Australia

84 IEEE TRANSACTIONS ON COGNITIVE AND DEVELOPMENTAL SYSTEMS, VOL. 8, NO. 2, JUNE 2016

Bootstrapping the Semantics of Tools: Affordance Analysis of Real World Objects on a Per-part Basis

Markus Schoeler and Florentin Wörgötter

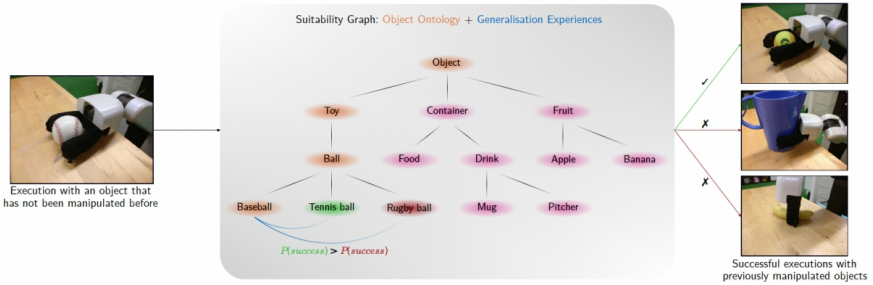
Motivation: Transferring Grasping Models Between Object Categories



A. Mitrevski, P. G. Plöger, and G. Lakemeyer, "A Hybrid Skill Parameterisation Model Combining Symbolic and Subsymbolic Elements for Introspective Robots," *Robotics and Autonomous Systems*, vol. 161, p. 104350:1–22, Mar. 2023. Available: https://doi.org/10.1016/j.robot.2022.104350

- **Given a collection of models** (e.g. for grasping different object) **and a previously unknown object, which of the existing models is the most suitable to reuse when executing a skill with the unknown object** (e.g. to grasp the object)?

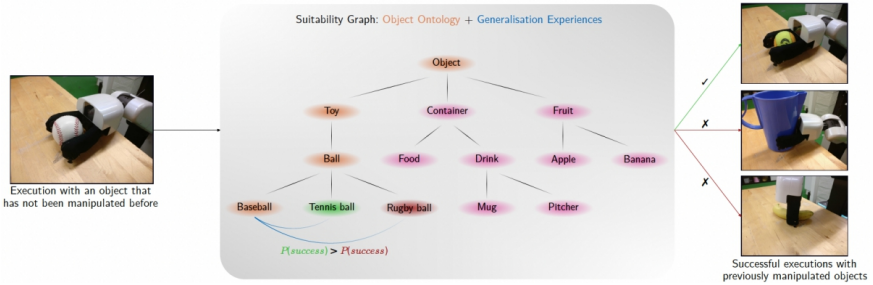
Motivation: Transferring Grasping Models Between Object Categories



A. Mitrevski, P. G. Plöger, and G. Lakemeyer, "A Hybrid Skill Parameterisation Model Combining Symbolic and Subsymbolic Elements for Introspective Robots," *Robotics and Autonomous Systems*, vol. 161, p. 104350:1–22, Mar. 2023. Available: https://doi.org/10.1016/j.robot.2022.104350

- **Given a collection of models** (e.g. for grasping different object) **and a previously unknown object, which of the existing models is the most suitable to reuse when executing a skill with the unknown object** (e.g. to grasp the object)?
- If the object were known, relations to other (known) categories could be used, but what if it is a completely new object that a robot has not interacted with before?

Motivation: Transferring Grasping Models Between Object Categories



A. Mitrevski, P. G. Plöger, and G. Lakemeyer, "A Hybrid Skill Parameterisation Model Combining Symbolic and Subsymbolic Elements for Introspective Robots," *Robotics and Autonomous Systems*, vol. 161, p. 104350:1–22, Mar. 2023. Available: https://doi.org/10.1016/j.robot.2022.104350

- **Given a collection of models** (e.g. for grasping different object) **and a previously unknown object, which of the existing models is the most suitable to reuse when executing a skill with the unknown object** (e.g. to grasp the object)?
- If the object were known, relations to other (known) categories could be used, but what if it is a completely new object that a robot has not interacted with before?
  - **Affordances can help**

Overview of Affordances

What is an Affordance?

James J. Gibson
From: The Ecological Approach to Visual Perception

Chapter 8
THE THEORY OF AFFORDANCES

- Humans usually **interact with objects in the environment in a structured way** — for instance, we hold a knife or a mug by its handle, or we grasp a glass by surrounding it

What is an Affordance?

James J. Gibson
From: The Ecological Approach to Visual Perception

Chapter 8
THE THEORY OF AFFORDANCES

- Humans usually **interact with objects in the environment in a structured way** — for instance, we hold a knife or a mug by its handle, or we grasp a glass by surrounding it
- An affordance is a **property associated with each object**, specifying **actions that can be performed on or with the object** (e.g. a glass affords pouring, a knife affords cutting)

What is an Affordance?

James J. Gibson
From: The Ecological Approach to Visual Perception

Chapter 8
THE THEORY OF AFFORDANCES

- Humans usually **interact with objects in the environment in a structured way** — for instance, we hold a knife or a mug by its handle, or we grasp a glass by surrounding it
- An affordance is a **property associated with each object**, specifying **actions that can be performed on or with the object** (e.g. a glass affords pouring, a knife affords cutting)
- Affordances can be thought of as **higher-level properties** that follow from lower-level properties, such as the size and shape of an object (e.g. a glass affords pouring because it is a tall, enclosed object made from a non-porous material)

What is an Affordance?

James J. Gibson
From: The Ecological Approach to Visual Perception

Chapter 8
THE THEORY OF AFFORDANCES

- Humans usually **interact with objects in the environment in a structured way** — for instance, we hold a knife or a mug by its handle, or we grasp a glass by surrounding it
- An affordance is a **property associated with each object**, specifying **actions that can be performed on or with the object** (e.g. a glass affords pouring, a knife affords cutting)
- Affordances can be thought of as **higher-level properties** that follow from lower-level properties, such as the size and shape of an object (e.g. a glass affords pouring because it is a tall, enclosed object made from a non-porous material)
- Different activities change the **way in which we use objects** — we cut using a knife by performing a slicing motion, but can open a bottle with a pushing motion on the cap

What is an Affordance?

James J. Gibson
From: The Ecological Approach to Visual Perception

Chapter 8
THE THEORY OF AFFORDANCES

- Humans usually **interact with objects in the environment in a structured way** — for instance, we hold a knife or a mug by its handle, or we grasp a glass by surrounding it
- An affordance is a **property associated with each object**, specifying **actions that can be performed on or with the object** (e.g. a glass affords pouring, a knife affords cutting)
- Affordances can be thought of as **higher-level properties** that follow from lower-level properties, such as the size and shape of an object (e.g. a glass affords pouring because it is a tall, enclosed object made from a non-porous material)
- Different activities change the **way in which we use objects** — we cut using a knife by performing a slicing motion, but can open a bottle with a pushing motion on the cap

"The affordances of the environment are what it offers the animal, what it provides or furnishes, either for good or ill." (Gibson 1979)

Affordances vs. Functionalities

It should be noted that **some authors make a distinction between the terms "affordance" and "functionality"**, such that there is a many-to-many relationship between the two:

Affordances vs. Functionalities

It should be noted that **some authors make a distinction between the terms "affordance" and "functionality"**, such that there is a many-to-many relationship between the two:

- **An affordance describes a property that can be perceived and that an object possesses inherently** (e.g. a knife handle is graspable)

Affordances vs. Functionalities

It should be noted that **some authors make a distinction between the terms "affordance" and "functionality"**, such that there is a many-to-many relationship between the two:

- **An affordance describes a property that can be perceived and that an object possesses inherently** (e.g. a knife handle is graspable)
- **A functionality relates an affordance to what a particular agent can do with an object** (e.g. a knife handle enables the agent to hold the knife)

Examples of Affordances

Affordance Label	References	Description	Examples
rollable	[5, 51, 81]	whether the object is rollable or not	roads, trolley
containment	[5, 97, 100, 126, 128]	indicates contain-ability of an object	pots
liquid-containment	[5]	indicates liquid contain-ability of an object	glasses, cups, mug
unstable	[5]	whether the object pose is stable after pushing	glass cups may be broken when pushing
stackable-onto	[5]	whether the object can be stacked	mugs, pots
sittable	[5, 61, 81, 119]	whether the object can be used to sit or not	chairs, desks
grasp	[51, 61, 81, 97, 100, 143]	defines the location of manipulation of flat tools	hammer, cups
cut	[97, 100, 126, 128, 143]	indicates cutting	knife, penknife, key
scoop	[97]	indicates curved surfaces tools	trowels, cookie scoop
pound	[97, 100]	indicates striking tools	hammerhead
support, place-on	[97, 100, 126]	indicates helpers or support an agent	flat tools (turners, spatulas), place-on (tables, desks), agent support (walls)
wrap-grasp	[97, 100]	detects the location of grasping	the outside of a cup)
pushable (forward, right, left)	[51, 61, 143]	whether the object is push-able	trolley, bike
liftable	[51, 61, 143]	whether the object can be lifted or no	liftable chairs
draggable, pushable backward	[51, 61]	whether the object can be dragged	desk, table
carryable	[51]	whether the object can be carried	light-weight pots, balls
traversable	[51]	whether the object can be traversed	road, grass
openable	[126, 143]	whether the object can be opened	fridge, room, microwave, book, box
pourable	[126]	whether the object is pour-able	mug
holdable	[126]	whether the object can be held	the outside of the mug
display, observe	[81, 100]	refers to display sources	TV, monitor screen
engine	[100]	refers to tools' engine parts	drill engine
hit	[100, 128]	refers to tools used to strike other objects.	racket head
obstruct	[81]	indicates the locations of obstructer	wall
break	[81]	indicates break-sensitive objects	glass cups
pinch-pull	[81]	indicates objects that pulled with a punch	knob
hook-pull	[81]	indicates objects that pulled with hooking up	handle
tip-push	[81]	indicates objects that perform actions after pushing	electricity buttons
warmth	[81]	indicates warmth objects	fireplaces
illumination	[81]	indicates light objects	lamps
dry	[81]	indicates objects that absorb water	towels
walk	[81, 119]	indicates places that allow walking	gardens
lyable	[119]	refers to free space that allow a person to lie down	bed
reachable	[119]	refers to reachable objects for picking up	bottle in the fridge
movable	[61]	refers to objects that can be moved around	small objects like balls, mugs

- Different affordance categories can be useful in different applications; thus, **most existing applications focus on a small subset of affordances**
- **Attempts to create general categorisations of objects based on their affordances have also been made** (as we will see in a few slides)
- The table on the right shows some affordances that have been used in various applications in the literature — most of these refer to **actions that can be directly done to objects** (e.g. an object can be pushed), while some describe **properties that hold due to those objects** (e.g. a lamp generates light)

Use of Affordances in Robotics

Affordances can be useful for various purposes, such as:

Providing context for performing various actions on objects

This makes it possible to use unknown objects, or to use known objects for non-standard purposes

Use of Affordances in Robotics

Affordances can be useful for various purposes, such as:

Providing context for performing various actions on objects

This makes it possible to use unknown objects, or to use known objects for non-standard purposes

Identifying possible actions that can and cannot be done with objects

For instance, affordances can provide important context for object grasping in terms of the part where an object should be held

Use of Affordances in Robotics

Affordances can be useful for various purposes, such as:

Providing context for performing various actions on objects

This makes it possible to use unknown objects, or to use known objects for non-standard purposes

Identifying possible actions that can and cannot be done with objects

For instance, affordances can provide important context for object grasping in terms of the part where an object should be held

Facilitating scene understanding

Affordances provide information about the activities that can be performed in a given scene and can thus guide exploration

Use of Affordances in Robotics

Affordances can be useful for various purposes, such as:

Providing context for performing various actions on objects

This makes it possible to use unknown objects, or to use known objects for non-standard purposes

Identifying possible actions that can and cannot be done with objects

For instance, affordances can provide important context for object grasping in terms of the part where an object should be held

Facilitating scene understanding

Affordances provide information about the activities that can be performed in a given scene and can thus guide exploration

Improving interaction in social scenarios

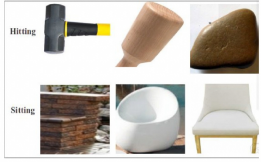
Even if certain actions can be performed on objects, the social context provides information about whether those actions are acceptable or desirable

Affordances in Cognitive Robotics

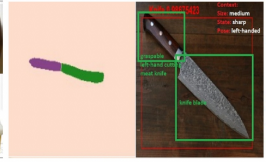
- The concept of affordances is **biologically inspired and is supported by various studies in psychology and developmental learning** — there is evidence that humans and other animals recognise and regularly use affordances
- The ability to recognise affordances is not acquired at birth, but observational studies of children provide evidence on how (and when) affordances are acquired in the infant years — the ability is **acquired through exploratory learning**
- Neuropsychological studies further provide evidence that **the recognition of objects is not a prerequisite for recognising affordances** — affordances allow us to interact with and use previously unseen objects in new ways

Affordance-Based Tasks

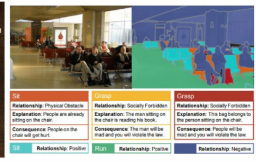
(a) Functionality understanding



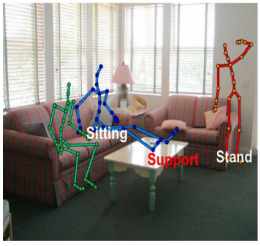
(b) Affordance categorization



(c) Affordance segmentation



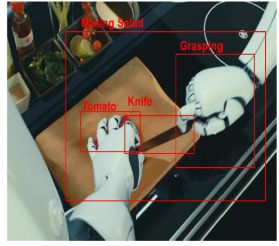
(d) Social affordances[19]



(e) Interactive Affordances

(f) Affordance detection

(g) Affordance reasoning



(h) Affordance-based activity recognition

Recognising Affordances

Recognising Affordances Preliminaries

- Visually, the recognition of affordances can be performed at different levels (in particular, at the level of individual objects or at the level of complete scenes); this gives rise to different types of affordance recognition tasks that may need to be performed

Recognising Affordances Preliminaries

MOHAMMED HASSANIN, University of New South Wales Canberra, Australia
SALMAN KHAN, Inception Institute of Artificial Intelligence, IIAT, UAE
MURAT TAHTALI, University of New South Wales Canberra, Australia

- Visually, the recognition of affordances can be performed at different levels (in particular, at the level of individual objects or at the level of complete scenes); this gives rise to different types of affordance recognition tasks that may need to be performed
- **Not all affordance recognition tasks are equally relevant for all applications** — grasping may require part-based recognition, while a coarse-grained recognition may suffice for scene understanding
  - In other words, the recognition of affordances at different levels is context-dependent

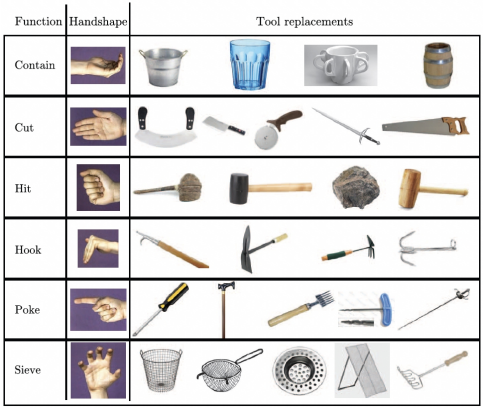
Recognising Affordances Preliminaries

- Visually, the recognition of affordances can be performed at different levels (in particular, at the level of individual objects or at the level of complete scenes); this gives rise to different types of affordance recognition tasks that may need to be performed
- **Not all affordance recognition tasks are equally relevant for all applications** — grasping may require part-based recognition, while a coarse-grained recognition may suffice for scene understanding
  - In other words, the recognition of affordances at different levels is context-dependent
- A cognitive robot would should be able to change the granularity of affordance recognition depending on the objectives of the overall task it is trying to achieve

Tool Affordance Ontology

Bootstrapping the Semantics of Tools: Affordance Analysis of Real World Objects on a Per-part Basis

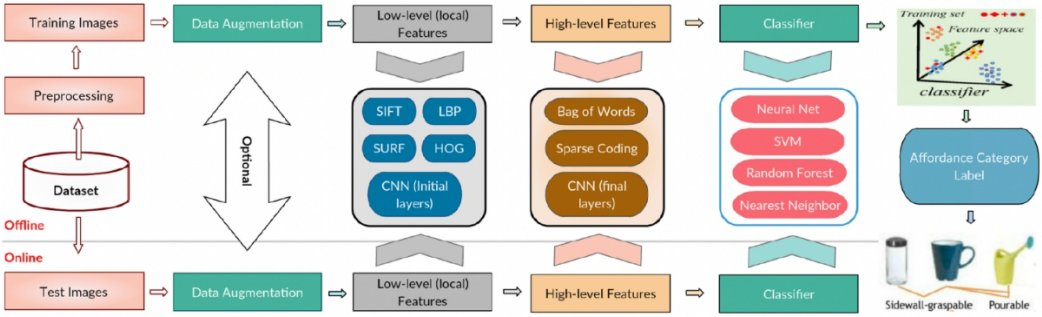
Markus Schoeler and Florentin Wörgötter



		Convex (Level 1)		Concave (Level 1)	
		Action (Level 3 / Level 4 / Level 5)	Tool	Action (Level 3 / Level 4 / Level 5)	Tool
Small AR (Level 2)	T	Paddle	Paddle	Sieve, lift	Sieve
		Hit	Rug beater	Sieve, shake	Sieve
		Spread	Butter knife, spatula	Shovel	Shovel
		Chop	Cleaver, axe, sword	Rake	Rake
		Cut	Knife, sword		
	R	Mix	Mixer	Empty	Shovel
		Paddle/mix parallel* circular	Blade(s) of an agitator		
		Paddle/mix perp.* circular	Blade(s)of a water mill		
		Grind parallel* circular	Grinding/millstone used flat		
		Cut/grind perp.* circular	Circular saw, angle grinder		
Medium AR (Level 2)	T	Push	Hammer	Fill	Cup, ladle
		Hit	Hammer		
		Grind	Grinder, pestle		
	R	Grind parallel* circular	Grinding/millstone used flat (fat)	Pour	Cup
		Grind perp.* circular	Grinding stone sed upright		
High AR (Level 2)	T	Push poke	Stick	Fill	Test tube
		Stab	Rapier, dagger		
		Draw, stir	Stick, pen		
		Push	Stick		
		Whip	Cane		
		Wipe, spread, hook	Stick, hook		
		Wipe, spread	Stick		
	R	Bore, drill	Drill	Pour	Test tube

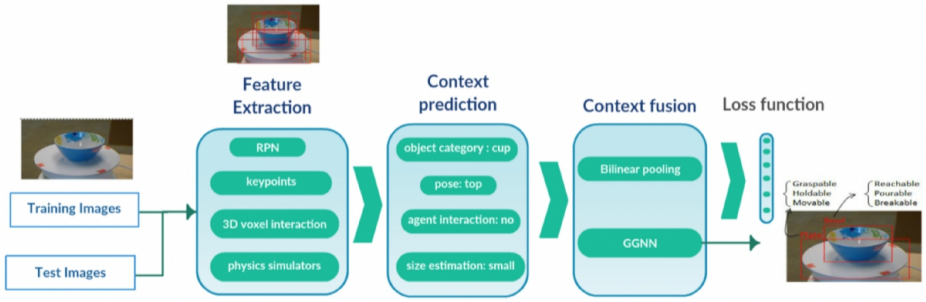
*perpendicular versus parallel refers to the orientation of the disk like tool relative to the target surface

Affordance Categorisation



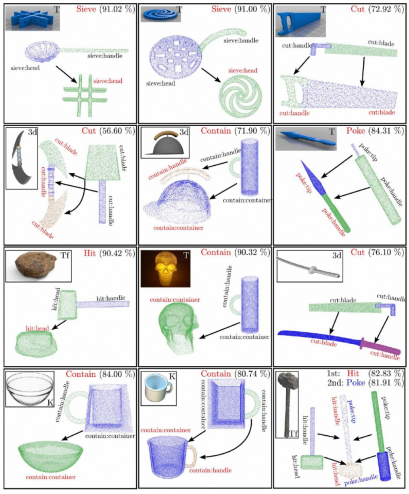
- Affordance categorisation is the problem of identifying an affordance in an object image (similar to an object recognition task)
- Traditionally modelled as a supervised learning problem — given a dataset of objects and their affordances, learn a classifier to recognise affordances

Affordance Detection



- Affordance detection can be seen as a joint problem of object detection and affordance categorisation (similar to object detection and recognition)
- Typically cast as a supervised learning problem — learn an object detection and affordance recognition model from a given dataset
- Such a model can also enable object recognition based on identified affordances

Affordance Segmentation



- Affordance segmentation is a more challenging problem than affordance detection and recognition — it is **concerned with identifying points belonging to different parts of an object**
- This requires recognising the affordance of the object as a whole, but also **detecting object parts and forming a structure of how parts are related to each other**
- Segmentation can be performed on 2D images or given 3D data (the approach illustrated on the right uses 3D point clouds)
  - 3D information can be directly consumed by a robot, e.g. for grasping

Affordances and Robot Learning

Learning (Based on) Affordances

- Information about affordances can be **used to learn execution models for various robot skills**, such as grasping or pushing objects

Learning (Based on) Affordances

- Information about affordances can be **used to learn execution models for various robot skills**, such as grasping or pushing objects
  - **Information about parts can be particularly relevant for handling objects correctly** (e.g. knives aren't typically grasped by the blade)

Learning (Based on) Affordances

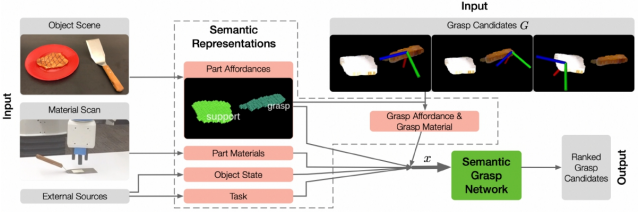
- Information about affordances can be **used to learn execution models for various robot skills**, such as grasping or pushing objects
  - **Information about parts can be particularly relevant for handling objects correctly** (e.g. knives aren't typically grasped by the blade)
  - **Affordance information can also be combined with task information**, such as the purpose for which an object is manipulated (e.g. a cup should be grasped at the handle for drinking, but can be grasped arbitrarily for transportation if it is empty)

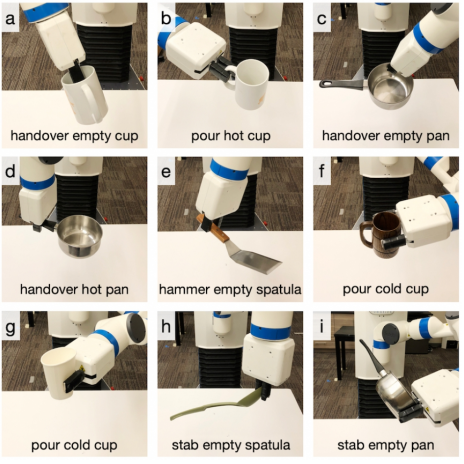
Learning (Based on) Affordances

- Information about affordances can be **used to learn execution models for various robot skills**, such as grasping or pushing objects
  - **Information about parts can be particularly relevant for handling objects correctly** (e.g. knives aren't typically grasped by the blade)
  - **Affordance information can also be combined with task information**, such as the purpose for which an object is manipulated (e.g. a cup should be grasped at the handle for drinking, but can be grasped arbitrarily for transportation if it is empty)
- Affordances are also **useful to consider when learning how different objects interact with each other** (e.g. tool-object interaction)

Grasp Learning Using Contextual and Affordance Information

- Combined with task and object state information, affordances can be used to learn a model that can generate appropriate grasping poses (as shown on the right)
- Earlier model of this type were based on probabilistic models (e.g. Bayesian networks); newer applications use deep neural networks as an underlying model

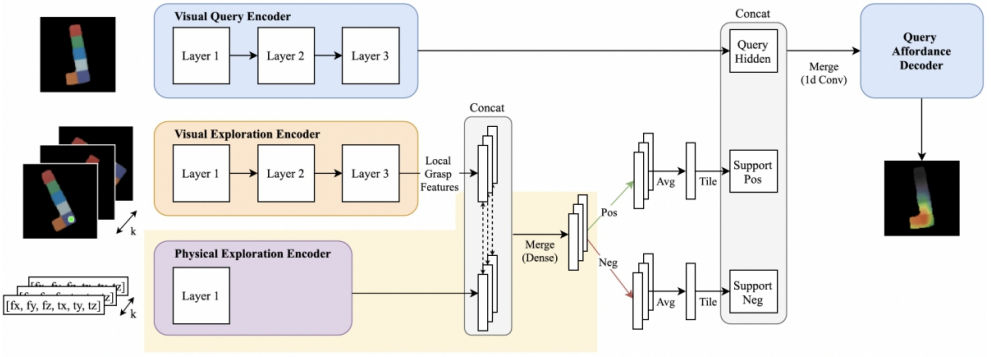




W. Liu et al., "CAGE: Context-Aware Grasping Engine," in *Proc. IEEE Int. Conf. Robotics and Automation (ICRA)*, 2020, pp. 2550–2556. Available: https://doi.org/10.1109/ICRA40945.2020.9197289

Multimodal Affordance-Based Learning

- Affordances are primarily recognised based on visual information, but this does not include all relevant object information, such as an object's mass, that can be relevant for execution
- Multimodal information can be combined to improve affordance predictions (for instance, visual information and by force / torque data, as shown below)

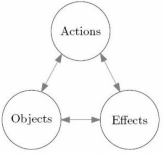


M. Veres et al., "Incorporating Object Intrinsic Features Within Deep Grasp Affordance Prediction," IEEE Robotics and Automation Letters (RAL), vol. 5, no. 4, pp. 6009–6016, Oct. 2020. Available: https://doi.org/10.1109/LRA.2020.3010444

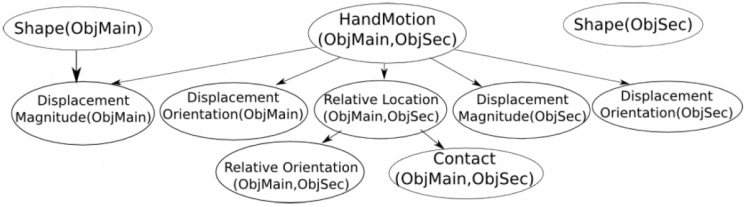
Learning Interaction Affordance Models

- Interaction models between objects can be learned by considering how the use of one object affects a second object — in other words, based on the action effects that can be observed when the objects interact with each other

- Below, an example of such a model is shown, which is represented by a Bayesian network that encodes the relationship between object shapes and relative displacements for different actions performed on the objects



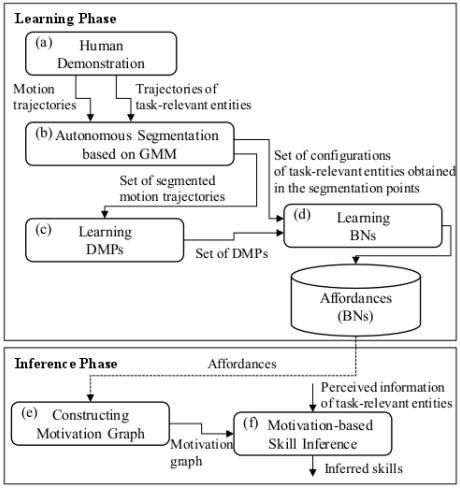
Inputs	Outputs	Function
(O, A)	E	Effect prediction
(O, E)	A	Action recognition/planning
(A, E)	O	Object recognition/selection



B. Moldovan et al., "Learning relational affordance models for robots in multi-object manipulation tasks," in *Proc. IEEE Int. Conf. Robotics and Automation (ICRA)*, 2012, pp. 4373–4378. Available: https://doi.org/10.1109/ICRA.2012.6225042

Skill Selection Based on Affordances

- Affordances are not useful only in the context of individual robot skills (e.g. grasping); they can also be used to inform the selection of skills themselves
- This can be achieved by creating a model that represents the observed effects of skills (similar to the object interaction model before) — in the case shown here, also a Bayesian network



S. H. Lee and I. H. Suh, "Skill learning and inference framework for skilligent robot," in *Proc. IEEE/RSJ Int. Conf. Intelligent Robots and Systems (IROS)*, 2013, pp. 108–115. Available: https://doi.org/10.1109/IROS.2013.6696340

Summary: Affordances for Cognitive Robots

- Affordances are an established concept from psychology that has been investigated in various studies of humans and other animals
- The role of affordances is to define actions that are enabled by or can be performed on objects in the environment
- Affordance recognition can be represented as a learning problem (different types of recognition problems can be defined)
- In robotics, affordances have been used to enhance the execution of concrete actions (e.g. object grasping), but also to represent the interaction between object and facilitate the selection of skills during execution
- In cognitive robotics, affordances are relevant because they can improve a robot's perceptual capabilities, facilitate exploratory learning, and guide the anticipation of action effects