Machine Common Sense Concept Paper

David Gunning

DARPA/I2O

arXiv '18

Presented by Koo hyeongseok
gudtjrdltka@korea.ac.kr
2021.01.14

All about Machine Commonsense

1) What is Commonsense and Why is it important?

2) Previous work in Machine Commonsense

3) Current work in Machine Commonsense

What is Commonsense and Why is it important?

1) What is Commonsense and Why is it important?

Common sense:

* Basic ability to perceive, understand, and judge things that are shared by nearly all people

* There is no need to state the obvious

Ex) 

"Can an elephant fit through the doorway?"

1) What is Commonsense and Why is it important?

However, because of obscure-but-pervasive nature of common sense,

Difficult to articulate and encode in machines

1) What is Commonsense and Why is it important?

Without general commonsense reasoning, machine reasoning is narrow and highly specialized

The absence of common sense prevents intelligent systems from

- Understanding their world
- Behaving reasonably in unforeseen situations
- Communicating naturally with people
- Learning from new experiences

Most significant challenge to achieve the more general, human-like AI systems

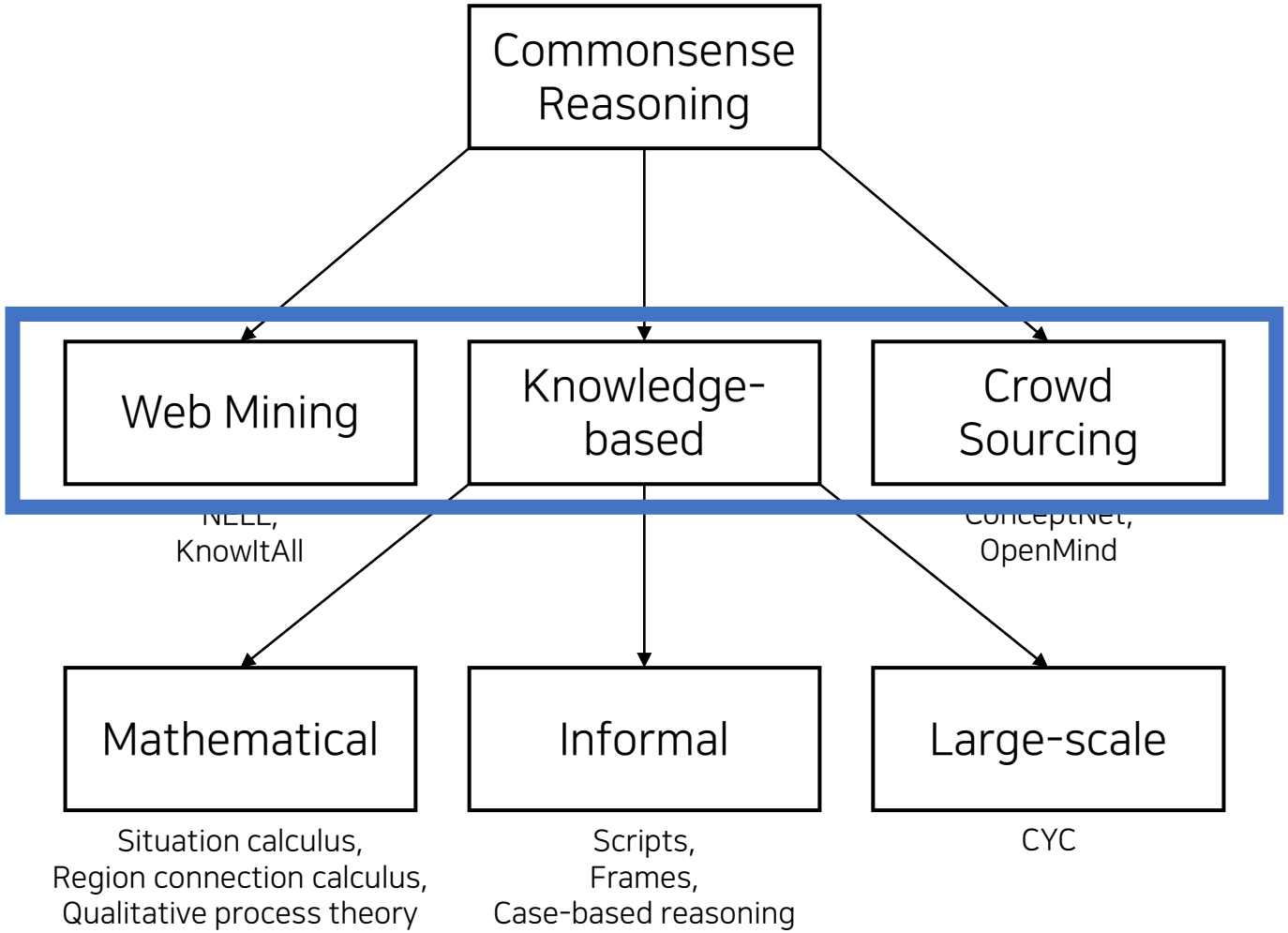
1) What is Commonsense and Why is it important?

Two strategies for developing machine commonsense

1. Construct computational models that mimic the core domains of child cognition
2. Construct a commonsense knowledge repository from reading the Web

Previous work in Machine Commonsense

2) Previous work in Machine Commonsense



[1] Davis, E., & Marcus, G. (2015). Commonsense reasoning and commonsense knowledge in artificial intelligence. Communications of the ACM, 58(9), 92-103.

2) Previous work in Machine Commonsense

1. Knowledge-based

1-1) Mathematical: Logic-based approaches

[1] Formal logic, [2] situation calculus, [3] naïve physics,
[4] default reasoning, ...

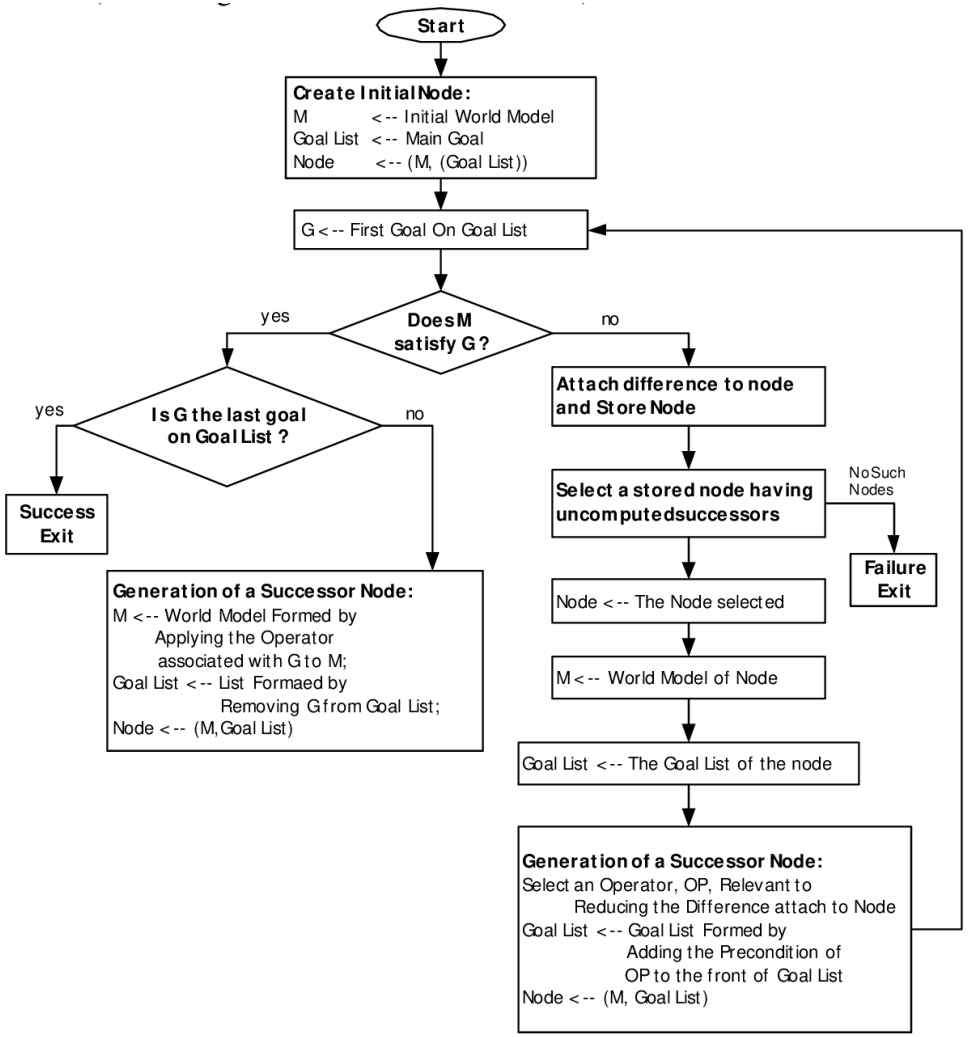


Fig. 2. Flow chart for STRIPS

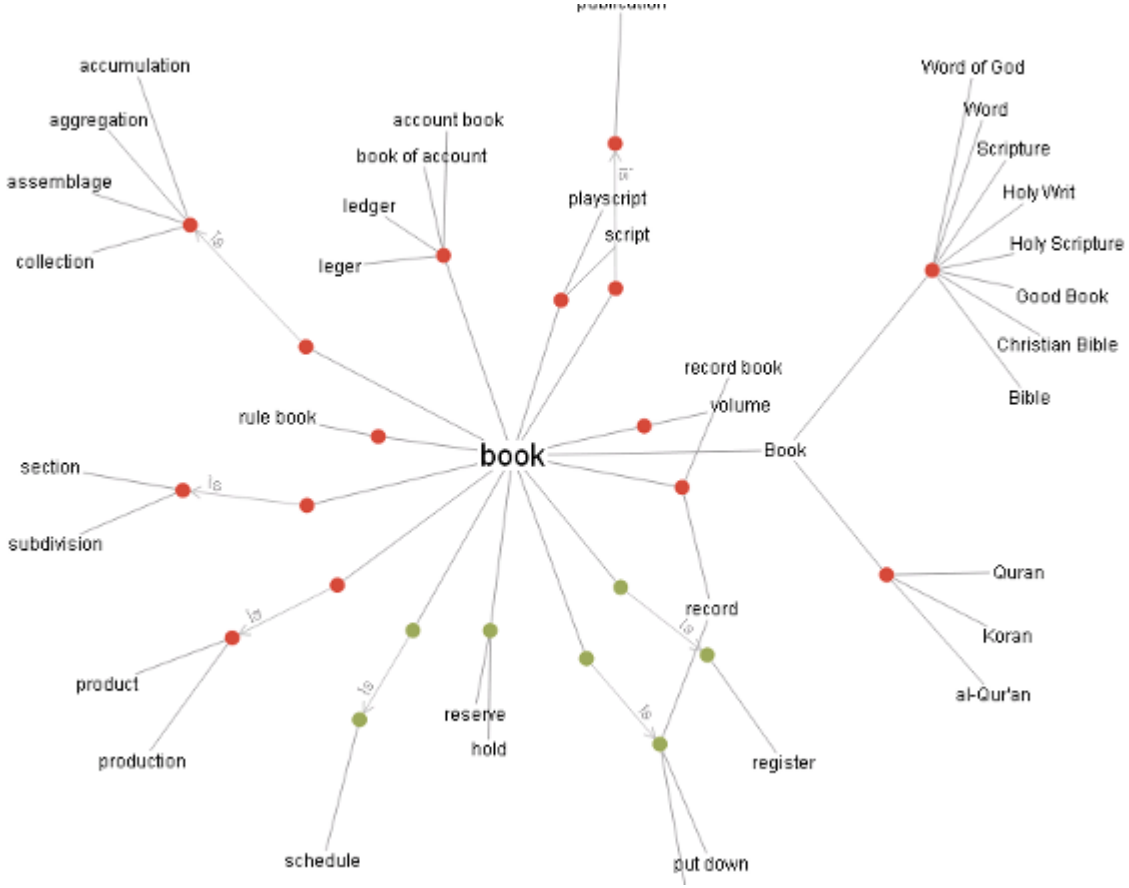
[1] McCarthy, J. (1960). Programs with common sense (pp. 300-307). RLE and MIT computation center.
[2] Fikes, R. E.. (1971). STRIPS: A new approach to the application of theorem proving to problem solving. Artificial intelligence, 2(3-4), 189-208.
[3] Hayes, P. J. (1978). The naive physics manifesto.
[4] Reiter, R. (1980). A logic for default reasoning. Artificial intelligence, 13(1-2), 81-132.

2) Previous work in Machine Commonsense

1. Knowledge-based

1-2) less formal knowledge-based approaches
(logic-based ontologies)

[1] frames, [2] scripts, [3] WordNet, [4] VerbNet, ...



[WordNet]

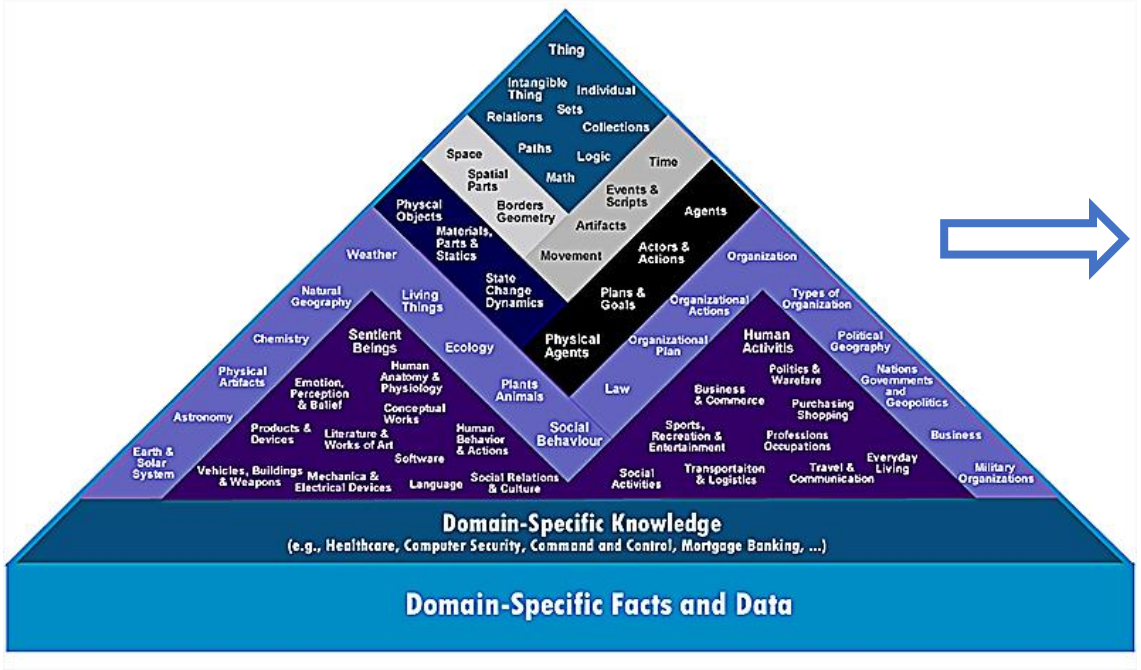
[1] Minsky, M. (1974). A framework for representing knowledge.
[2] Schank, R. C., & Abelson, R. P. (1975, September). Scripts, plans, and knowledge. In IJCAI (pp.151-157).
[3] Miller, G. (1998). WordNet: An electronic lexical database. MIT press.
[4] Schuler, K. K. (2005). VerbNet: A broad-coverage, comprehensive verb lexicon.

2) Previous work in Machine Commonsense

1. Knowledge-based

1-3) Large-scale logic-based system

Cyc[1]: a 35-year effort to codify common sense into an integrated, logic-based system



- It covers large areas of commonsense knowledge
- It integrates sophisticated, logic-based reasoning techniques.

<u>However,</u>

- Never quite match with human concepts & natural language queries
- Need to be tailored and refined to fit specific applications

[1] Lenat, D. B. (1995). CYC: A large-scale investment in knowledge infrastructure. Communications of the ACM, 38(11), 33-38.

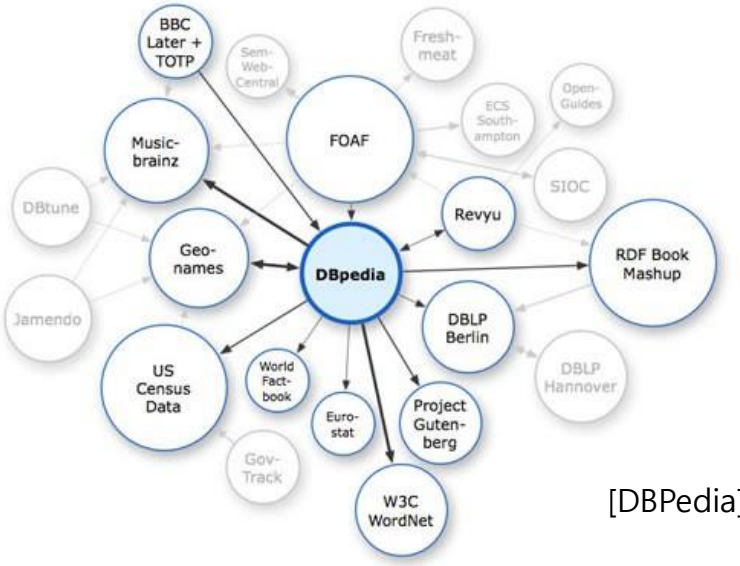
2) Previous work in Machine Commonsense

2. Web Mining, Crowd Sourcing

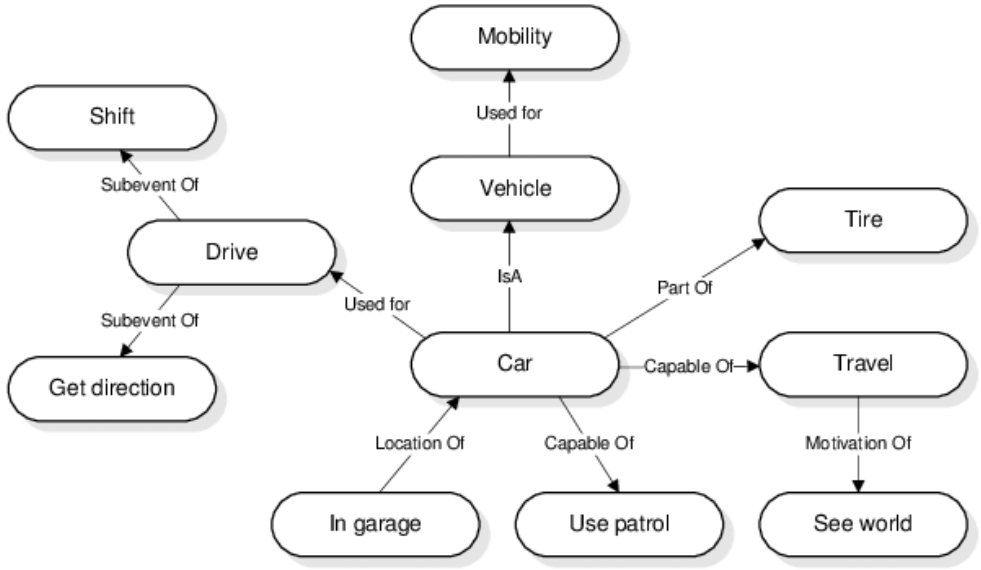
: extract and collect commonsense knowledge from the Web

2-1) Web Mining
: from the entire Web (e.g., KnowItAll [1]), from a subset of the Web such as Wikipedia (e.g., DBPedia [2]).

2-2) Crowd Sourcing
: from the general public via the Web, such as OpenMind [3] and ConceptNet [4].



[DBPedia]



[ConceptNet]

[1] Etzioni, O., Web-scale information extraction in knowitall:(preliminary results). In Proceedings of the 13th international conference on World Wide Web (pp. 100-110). ACM.
[2] Auer, S., Bizer, C., Kobilarov, G., Lehmann, J., Cyganiak, R., & Ives, Z. (2007). Dbpedia: A nucleus for a web of open data. In The semantic web (pp. 722-735). Springer, Berlin, Heidelberg.
[3] Singh, P., Open Mind Common Sense: Knowledge acquisition from the general public. In OTM Confederated International Conferences" On the Move to Meaningful Internet Systems" (pp. 1223-1237). Springer, Berlin, Heidelberg.
[4] Liu, H., & Singh, P. (2004). ConceptNet—a practical commonsense reasoning tool-kit. BT technology journal, 22(4), 211-226.

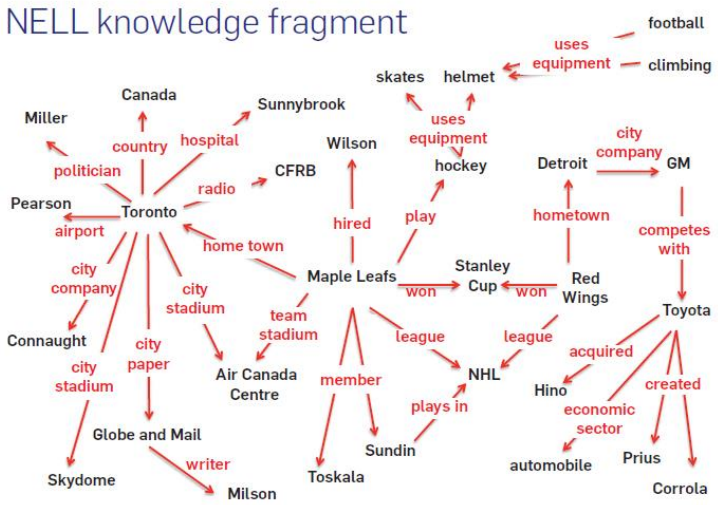
2) Previous work in Machine Commonsense

2. Web Mining, Crowd Sourcing

2-3) Combining Web Mining & Crowd Sourcing

Nell(Never Ending Language Learning)[1]

- An initial seed ontology defining hundreds of categories and relations that NELL is expected to read about
- Running continuously to extract new instances of categories and relations
- Using crowdsourcing to provide feedback from humans in order to improve the quality of its extractions



Make significant progress, but,

- shallow semantic representations suffer from ambiguities and inconsistencies
- Lack sufficient semantic understanding to enable reasoning beyond simple answer lookup

[1] Carlson, A., Betteridge, J., Kisiel, B., Settles, B., Hruschka Jr, E. R., & Mitchell, T. M. (2010, July). Toward an architecture for never-ending language learning. In AAAI (Vol. 5, p. 3).
[2] Mitchell, T., Cohen, W., Hruschka, E., Talukdar, P., Yang, B., Betteridge, J., & Krishnamurthy, J.(2018). Never-ending learning. Communications of the ACM, 61(5), 103-115.

2) Previous work in Machine Commonsense

Need flexible, perceptually grounded concept representations!

Human Cognition - Theory of Grounded Cognition

Ex. “door”

⇒ Perceptual-motor simulation: Can imagine opening the door using your arm

Ex. “when one door closes, another opens”

⇒ Perceptual-motor experience: Understand the meaning of that abstract idea

Perceptually grounded representations is critical for making progress on machine common sense

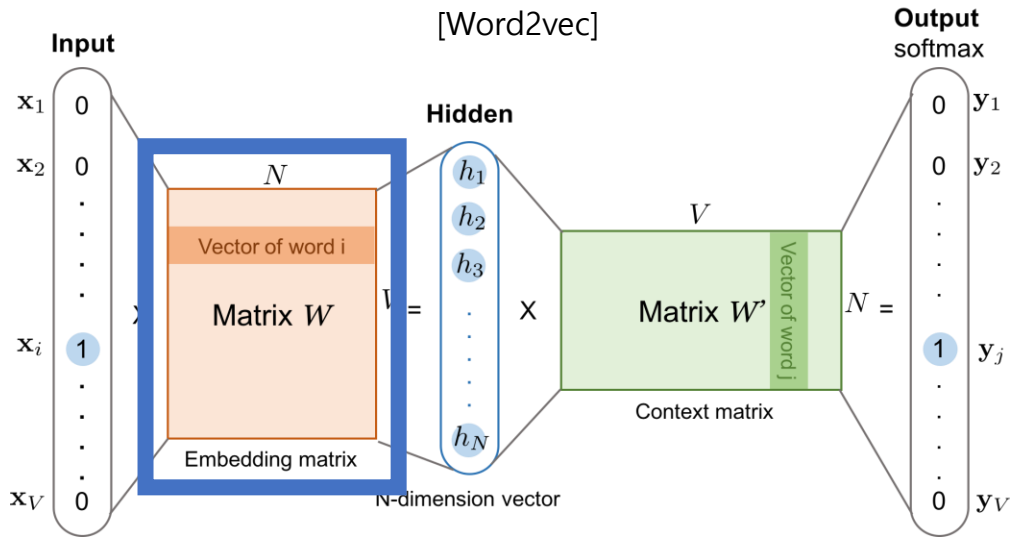
Current work in Machine Commonsense

3) Current work in Machine Commonsense

1. Grounded Representations

Use of vector-based embeddings - word2vec[1][2]

- Embeddings are widely used to learn representations.
- Representations are from language, images, video etc.



However,

- These representations are not perfect and have limitations
- Trying to discover new techniques to effectively compose, simulate, and reason with these representations

<u>These new representations are foundations for learning human-like common sense concepts.</u>

[1] Mikolov, T., Distributed representations of words and phrases and their compositionality. In Advances in neural information processing systems (pp. 3111-3119).
[2] Le, Q., & Mikolov, T. (2014, January). Distributed representations of sentences and documents. In International Conference on Machine Learning (pp. 1188-1196)..

3) Current work in Machine Commonsense

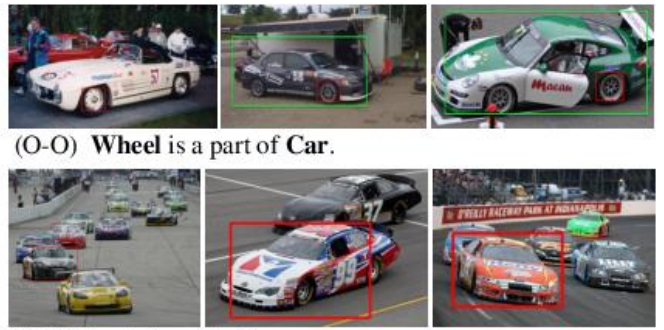
2. Learning Commonsense Knowledge from the Web

: Learning commonsense knowledge from images and language on the Web.

(a) Objects (w/Bounding Boxes and Visual Subcategories)

(b) Scenes (c) Attributes

Visual Instances Labeled by NEIL



(O-O) **Wheel** is a part of **Car**.

(S-O) **Car** is found in **Raceway**.

(O-O) **Corolla** is a kind of/looks similar to **Car**.

(S-O) **Pyramid** is found in **Egypt**.

(O-A) **Wheel** is/has **Round shape**.
(S-A) **Alley** is/has **Narrow**.
(S-A) **Bamboo forest** is/has **Vertical lines**.
(O-A) **Sunflower** is/has **Yellow**.

Relationships Extracted by NEIL

[1] Chen, X., Shrivastava, A., & Gupta, A. (2013). Neil: Extracting visual knowledge from web data. In Proceedings of the IEEE International Conference on Computer Vision (pp. 1409-1416).

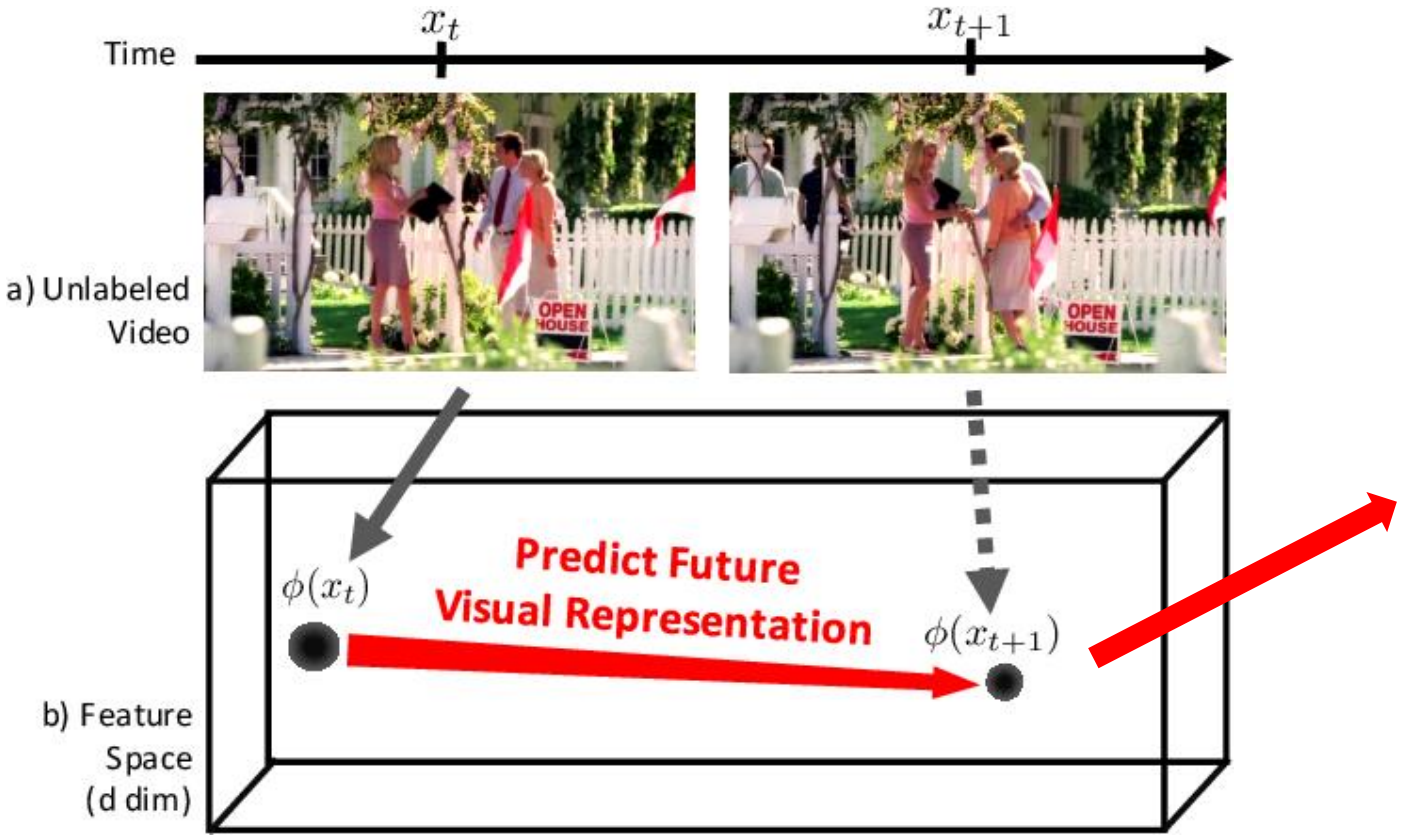
3) Current work in Machine Commonsense

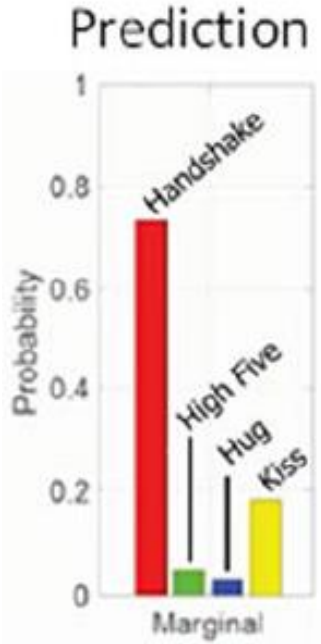
3. Learning Predictive Models from Experience

: Learning predictive models of commonsense phenomenon from videos and simulations

self-supervised techniques

: Predict the next event in an unlabeled video sequence





[1] Vondrick, C., Pirsiavash, H., & Torralba, A. (2016). Anticipating visual representations from unlabeled video. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 98-106).

3) Current work in Machine Commonsense

4. Understanding and Modeling Childhood Cognition

: Cognitive development can provide guidance for building intelligent machines that think and learn like children

Six core domains of children's knowledge

Domain	Description
Objects	supports reasoning about objects and the laws of physics that govern them
Agents	supports reasoning about agents that act autonomously to pursue goals
Places	supports navigation and spatial reasoning around an environment
Number	supports reasoning about quantity and how many things are present
Forms	supports representation of shapes and their affordances
Social Beings	supports reasoning about Theory of Mind and social interactions

These core domains serve as the fundamental building blocks of human intelligence and common sense

objects (intuitive physics), agents (intentional actors), and places (spatial navigation)

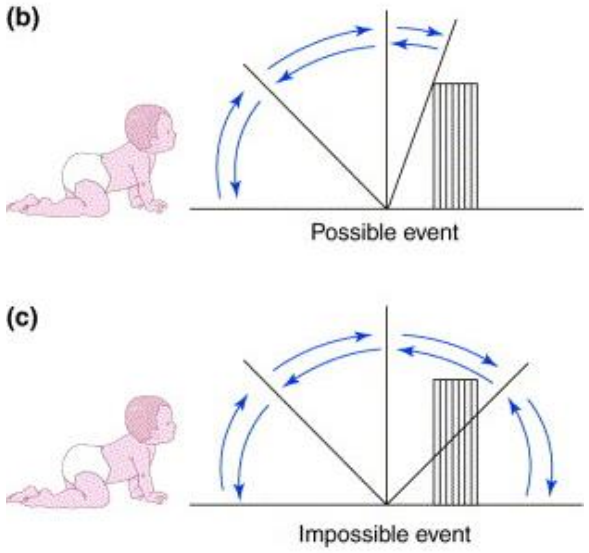
[1] Chen, X., Shrivastava, A., & Gupta, A. (2013). Neil: Extracting visual knowledge from web data. In Proceedings of the IEEE International Conference on Computer Vision (pp. 1409-1416).

3) Current work in Machine Commonsense

4. Understanding and Modeling Childhood Cognition

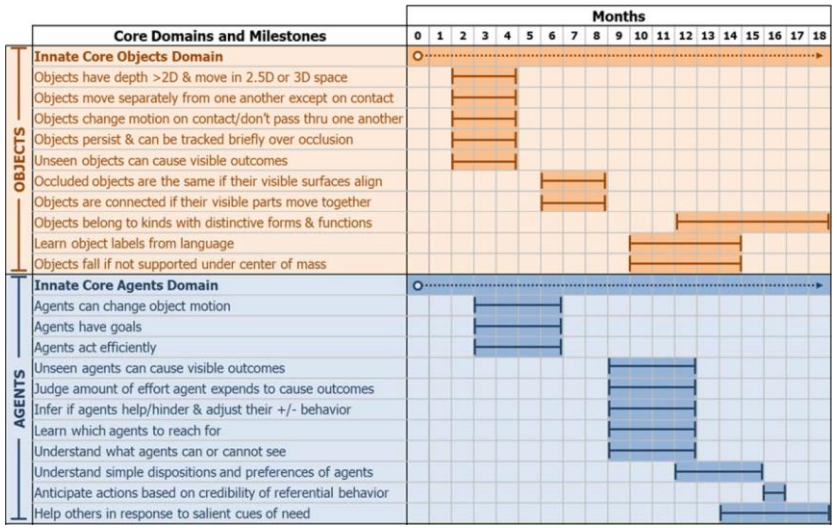
Violation of Expectation (VOE):

-Observing a child's surprise
–Primary means of studying child cognitive abilities



Key stages of the developmental sequence:

-The domains of objects, agents, and places
-An excellent set of target milestones for AI researchers

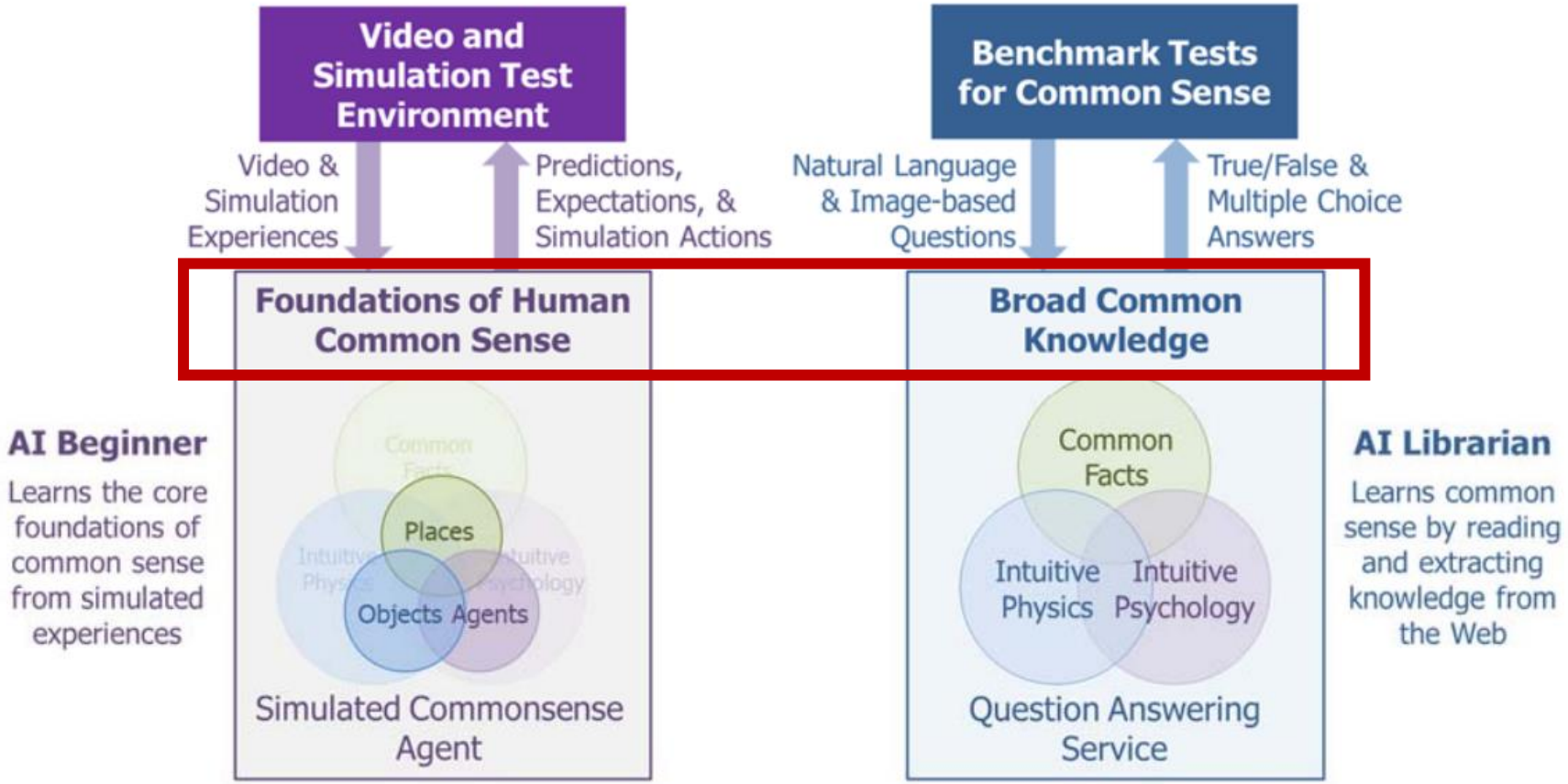


AI researchers have begun to use these results to create computational models of child cognition

Ex) Computational models of intuitive physics that mimic child cognition(MIT)

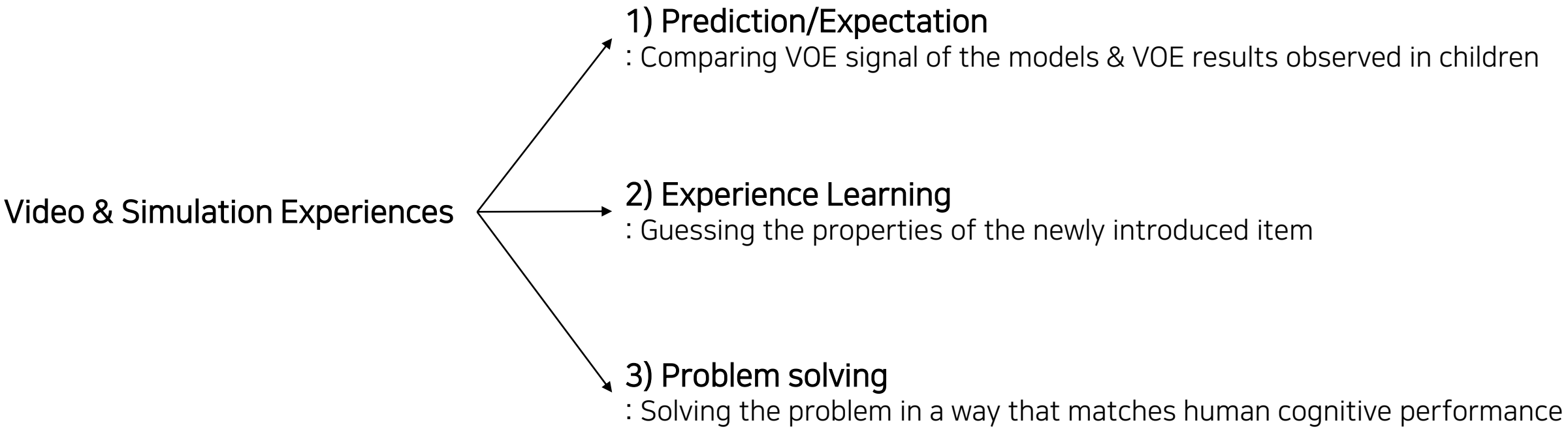
How to evaluate machine commonsense?

How to evaluate machine commonsense?



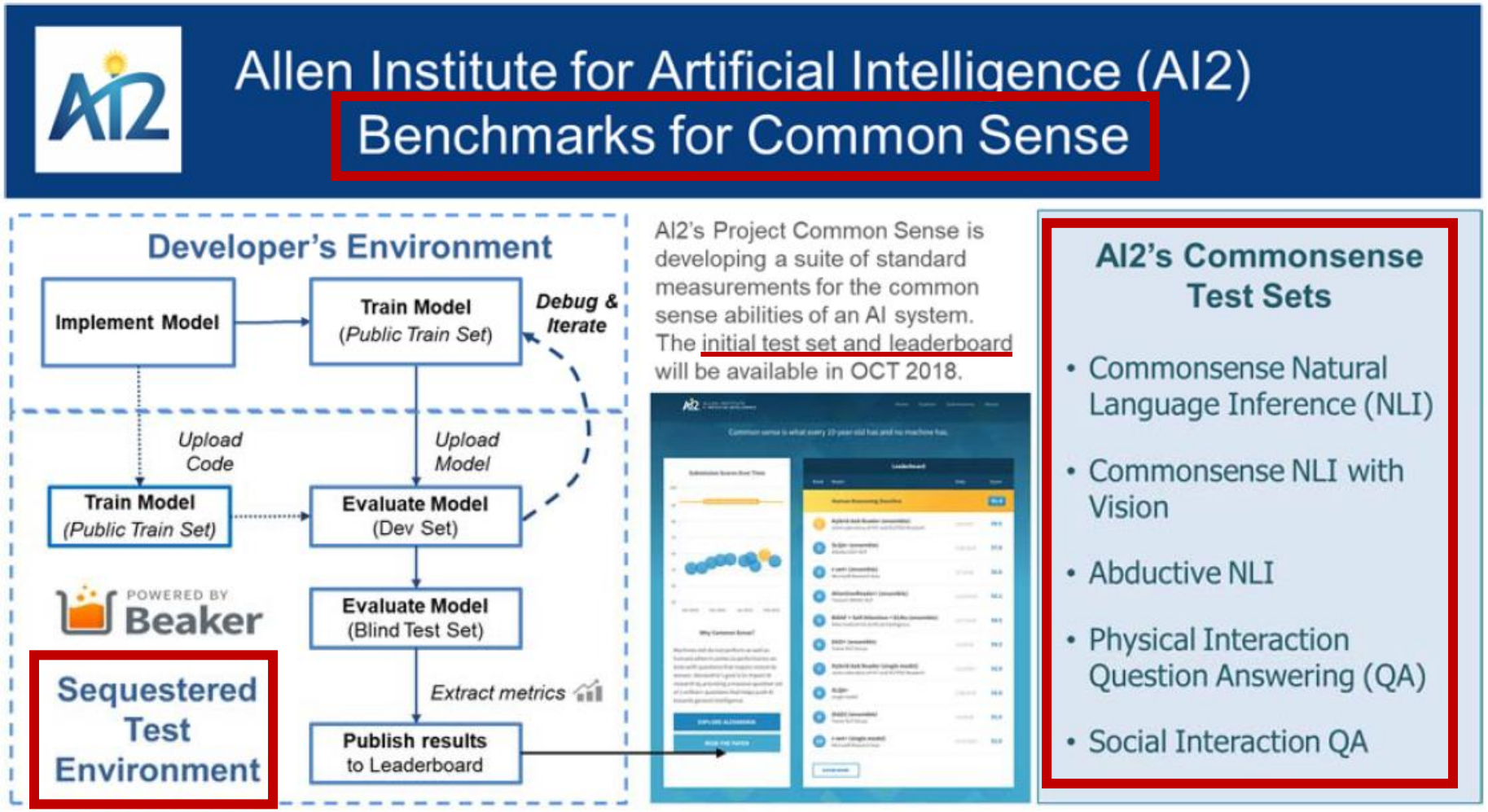
How to evaluate machine commonsense?

1) Foundations of Human Common Sense



How to evaluate machine commonsense?

2) Broad common knowledge
: capable of answering natural language and image-based questions about commonsense phenomena



감사합니다