From Relations to Random Variables and Features – Lecture 17.1

Topics:

- Reconciling relations and random variables / features
- From knowledge graphs to random variables
- More general relationships

"The mind is a neural computer, fitted by natural selection with combinatorial algorithms for causal and probabilistic reasoning about plants, animals, objects, and people. It is driven by goal states that served biological fitness in ancestral environments, such as food, sex, safety, parenthood, friendship, status and knowledge."

...

"In a universe with any regularities at all, decisions informed about the past are better than decisions made at random. That has always been true, and we would expect organisms, especially informavores such as humans, to have evolved acute intuitions about probability. The founders of probability, like the founders of logic, assumed they were just formalizing common sense."

Steven Pinker, How the Mind Works, 1997, pp. 524, 343.

"The mind is a neural computer, fitted by natural selection with combinatorial algorithms for causal and probabilistic reasoning about plants, animals, objects, and people. It is driven by goal states that served biological fitness in ancestral environments, such as food, sex, safety, parenthood, friendship, status and knowledge."

. . .

"In a universe with any regularities at all, decisions informed about the past are better than decisions made at random. That has always been true, and we would expect organisms, especially informavores such as humans, to have evolved acute intuitions about probability. The founders of probability, like the founders of logic, assumed they were just formalizing common sense."

Steven Pinker, How the Mind Works, 1997, pp. 524, 343.

Real World

What is the real world made of?

A Features or random variables

B Words, pixels, phonemes . . .

C Entities and events (e.g., plants, people, diseases, lectures, university course)

D Huh? There is a real world?

Reconciling Logic and Probability

- How to reconcile
  - Features and random variables
  - Entities/things/objects/events

Reconciling Logic and Probability

- How to reconcile
  - Features and random variables
  - Entities/things/objects/events
- Entities are not features or random variables.

Reconciling Logic and Probability

- How to reconcile
  - Features and random variables
  - Entities/things/objects/events
- Entities are not features or random variables.
- It makes no sense to talk about the probability of a person.

Reconciling Logic and Probability

- How to reconcile
  - Features and random variables
  - Entities/things/objects/events
- Entities are not features or random variables.
- It makes no sense to talk about the probability of a person. Compare:
  - The probability of Shakira
  - The probability that Shakira will record a song with Drake next year.

Reconciling Logic and Probability

- How to reconcile
  - Features and random variables
  - Entities/things/objects/events
- Entities are not features or random variables.
- It makes no sense to talk about the probability of a person. Compare:
  - The probability of Shakira
  - The probability that Shakira will record a song with Drake next year.
- The word "variable" has different meanings in probability and logic.
  - In logic a variable denotes an entity.
  - In probability a variable denotes a function over possible worlds (that we may be uncertain of the value of).

Statistical Relational AI

Statistical relational AI is about making predictions about properties of entities and relations among entities.

Statistical Relational AI

Statistical relational AI is about making predictions about properties of entities and relations among entities.

Tasks:

- Predicting attributes of an entity based on its other attributes and attributes of related entities.

Statistical Relational AI

Statistical relational AI is about making predictions about properties of entities and relations among entities.

Tasks:

- Predicting attributes of an entity based on its other attributes and attributes of related entities.
  E.g, predicting age of a person from their postal code and what movies that have rated.

Statistical Relational AI

Statistical relational AI is about making predictions about properties of entities and relations among entities.

Tasks:

- Predicting attributes of an entity based on its other attributes and attributes of related entities.
  E.g, predicting age of a person from their postal code and what movies that have rated.
- Predicting relations among multiple entities based on properties and relations of the entities involved.

Statistical Relational AI

Statistical relational AI is about making predictions about properties of entities and relations among entities.

Tasks:

- Predicting attributes of an entity based on its other attributes and attributes of related entities.
  E.g, predicting age of a person from their postal code and what movies that have rated.
- Predicting relations among multiple entities based on properties and relations of the entities involved.
  E.g, predict what rating a user will give a movie

Statistical Relational AI

Statistical relational AI is about making predictions about properties of entities and relations among entities.

Tasks:

- Predicting attributes of an entity based on its other attributes and attributes of related entities.
  E.g, predicting age of a person from their postal code and what movies that have rated.
- Predicting relations among multiple entities based on properties and relations of the entities involved.
  E.g, predict what rating a user will give a movie
- Predicting identity, whether descriptions denote the same entity

Statistical Relational AI

Statistical relational AI is about making predictions about properties of entities and relations among entities.
Tasks:

- Predicting attributes of an entity based on its other attributes and attributes of related entities.
  E.g, predicting age of a person from their postal code and what movies that have rated.
- Predicting relations among multiple entities based on properties and relations of the entities involved.
  E.g, predict what rating a user will give a movie
- Predicting identity, whether descriptions denote the same entity
  E.g., which citations refer to the same paper

Statistical Relational AI

Statistical relational AI is about making predictions about properties of entities and relations among entities.

Tasks:

- Predicting attributes of an entity based on its other attributes and attributes of related entities.
  E.g, predicting age of a person from their postal code and what movies that have rated.
- Predicting relations among multiple entities based on properties and relations of the entities involved.
  E.g, predict what rating a user will give a movie
- Predicting identity, whether descriptions denote the same entity
  E.g., which citations refer to the same paper
- Predicting existence, whether an entity exists that fits a description

Statistical Relational AI

Statistical relational AI is about making predictions about properties of entities and relations among entities.

Tasks:

- Predicting attributes of an entity based on its other attributes and attributes of related entities.
  E.g, predicting age of a person from their postal code and what movies that have rated.
- Predicting relations among multiple entities based on properties and relations of the entities involved.
  E.g, predict what rating a user will give a movie
- Predicting identity, whether descriptions denote the same entity
  E.g., which citations refer to the same paper
- Predicting existence, whether an entity exists that fits a description
  E.g., whether there is a person in a particular room

From Knowledge Graphs to Random Variables

- A knowledge graph consists of triples of form $\langle s, r, o \rangle$ with subject s, relation (verb) r, and object o. Also written $r(s, o)$.

From Knowledge Graphs to Random Variables

- A knowledge graph consists of triples of form $\langle s, r, o \rangle$ with subject s, relation (verb) r, and object o. Also written $r(s, o)$.
- r is functional means there is at most one object for a subject: (x, r, y_1) and (x, r, y_2) implies $y_1 = y_2$.

From Knowledge Graphs to Random Variables

- A knowledge graph consists of triples of form $\langle s, r, o \rangle$ with subject s, relation (verb) r, and object o. Also written $r(s, o)$.
- r is functional means there is at most one object for a subject: (x, r, y_1) and (x, r, y_2) implies $y_1 = y_2$.
  E.g., *height* (on January 1, 2026), *date-of-birth*, *birth-mother*

From Knowledge Graphs to Random Variables

- A knowledge graph consists of triples of form $\langle s, r, o \rangle$ with subject s, relation (verb) r, and object o. Also written $r(s, o)$.
- r is functional means there is at most one object for a subject: (x, r, y_1) and (x, r, y_2) implies $y_1 = y_2$.
  E.g., *height* (on January 1, 2026), *date-of-birth*, *birth-mother*
- For functional relations, there is a random variable for each subject-relation pair.

From Knowledge Graphs to Random Variables

- A knowledge graph consists of triples of form $\langle s, r, o \rangle$ with subject s, relation (verb) r, and object o. Also written $r(s, o)$.
- r is functional means there is at most one object for a subject: (x, r, y_1) and (x, r, y_2) implies $y_1 = y_2$.
  E.g., *height* (on January 1, 2026), *date-of-birth*, *birth-mother*
- For functional relations, there is a random variable for each subject-relation pair.
  E.g., for each person, *birth mother* has a distribution over people

From Knowledge Graphs to Random Variables

- A knowledge graph consists of triples of form $\langle s, r, o \rangle$ with subject s, relation (verb) r, and object o. Also written $r(s, o)$.
- r is functional means there is at most one object for a subject: (x, r, y_1) and (x, r, y_2) implies $y_1 = y_2$.
  E.g., *height* (on January 1, 2026), *date-of-birth*, *birth-mother*
- For functional relations, there is a random variable for each subject-relation pair.
  E.g., for each person, *birth mother* has a distribution over people
  The domain of the random variable is the range of the function.

From Knowledge Graphs to Random Variables

- A knowledge graph consists of triples of form $\langle s, r, o \rangle$ with subject s, relation (verb) r, and object o. Also written $r(s, o)$.
- r is functional means there is at most one object for a subject: (x, r, y_1) and (x, r, y_2) implies $y_1 = y_2$.
  E.g., *height* (on January 1, 2026), *date-of-birth*, *birth-mother*
- For functional relations, there is a random variable for each subject-relation pair.
  E.g., for each person, *birth mother* has a distribution over people
  The domain of the random variable is the range of the function.
- For non-functional properties, there is a Boolean random variable for each subject–relation–object triple.

From Knowledge Graphs to Random Variables

- A knowledge graph consists of triples of form $\langle s, r, o\rangle$ with subject s, relation (verb) r, and object o. Also written $r(s, o)$.
- r is functional means there is at most one object for a subject: (x, r, y_1) and (x, r, y_2) implies $y_1 = y_2$.
  E.g., *height* (on January 1, 2026), *date-of-birth*, *birth-mother*
- For functional relations, there is a random variable for each subject-relation pair.
  E.g., for each person, *birth mother* has a distribution over people
  The domain of the random variable is the range of the function.
- For non-functional properties, there is a Boolean random variable for each subject–relation–object triple.
  E.g., relation *has-streamed* relation between person and a musical artist

From Knowledge Graphs to Random Variables

For more general relationships $r(X_1, \ldots, X_k)$:

- If one argument, say X_k, is a function of the other arguments, there is a random variable for each tuple $r(e_1, \ldots, e_{k-1})$ where the domain of the random variable is the set of values that X_k can take.

From Knowledge Graphs to Random Variables

For more general relationships $r(X_1, \ldots, X_k)$:

- If one argument, say X_k, is a function of the other arguments, there is a random variable for each tuple $r(e_1, \ldots, e_{k-1})$ where the domain of the random variable is the set of values that X_k can take.
  E.g. the relation $rated(U, M, R)$ – user U gave movie M a rating of R (from 1 to 5) – gives a random variable for each user–movie pair with domain the set of possible ratings, $\{1, 2, 3, 4, 5\}$.

From Knowledge Graphs to Random Variables

For more general relationships $r(X_1, \ldots, X_k)$:

- If one argument, say X_k, is a function of the other arguments, there is a random variable for each tuple $r(e_1, \ldots, e_{k-1})$ where the domain of the random variable is the set of values that X_k can take.
  E.g. the relation $rated(U, M, R)$ – user U gave movie M a rating of R (from 1 to 5) – gives a random variable for each user–movie pair with domain the set of possible ratings, $\{1, 2, 3, 4, 5\}$.
- Otherwise, there is a Boolean random variable for each tuple $r(e_1, \ldots, e_k)$.

From Knowledge Graphs to Random Variables

For more general relationships $r(X_1, \ldots, X_k)$:

- If one argument, say X_k, is a function of the other arguments, there is a random variable for each tuple $r(e_1, \ldots, e_{k-1})$ where the domain of the random variable is the set of values that X_k can take.
  E.g. the relation $rated(U, M, R)$ – user U gave movie M a rating of R (from 1 to 5) – gives a random variable for each user–movie pair with domain the set of possible ratings, $\{1, 2, 3, 4, 5\}$.
- Otherwise, there is a Boolean random variable for each tuple $r(e_1, \ldots, e_k)$.

The functional case is treated as a relation of $k - 1$ arguments, with a non-Boolean prediction.