Banff International Research Station
for Mathematical Innovation and Discovery

Probabilistic Models of Cognitive Development
May 25-29, 2009

MEALS

*Breakfast (Buffet): 7:00 – 9:30 am, Sally Borden Building, Monday – Friday
*Lunch (Buffet): 11:30 am – 1:30 pm, Sally Borden Building, Monday – Friday
*Dinner (Buffet): 5:30 – 7:30 pm, Sally Borden Building, Sunday – Thursday
Coffee Breaks: As per daily schedule, 2nd floor lounge, Corbett Hall
***Please remember to scan your meal card at the host/hostess station in the dining room for each meal.**

MEETING ROOMS

All lectures will be held in Max Bell 159 (Max Bell Building accessible by walkway on 2nd floor of Corbett Hall). LCD projector, overhead projectors and blackboards are available for presentations. *Please note that the meeting space designated for BIRS is the lower level of Max Bell, Rooms 155-159. Please respect that all other space has been contracted to other Banff Centre guests, including any Food and Beverages in those areas.*

SCHEDULE

Sunday

4:00	Check-in begins (Front Desk – Professional Development Centre - open 24 hours)
5:30-7:30	Buffet Dinner
8:00	Informal gathering in 2nd floor lounge, Corbett Hall Beverages and small assortment of snacks available on a cash honour-system.

Monday

7:00-8:45	Breakfast
8:45-9:00	Introduction and Welcome to BIRS by BIRS Station Manager (in Max Bell 159)
9:00-9:35	Introduction by Fei Xu (from perspective of development) and Tom Griffiths (from perspective of computation)
9:35-10:15	Probabilistic Reasoning in Infants and Children (Bonatti, Girotto, Xu) *Predicting the future without knowing the past: Infants' sensitivity to probabilities* Luca Bonatti
10:15-10:40	Coffee break, 2nd floor lounge, Corbett Hall
10:40-12:00	Probabilistic Reasoning in Infants and Children (continued) *Children's probabilistic reasoning about combinatorial possibilities* Vittorio Girotto *Rational statistical inference in infants* Fei Xu
12:00-12:30	Discussion (Moderator Tom Griffiths)

12:30-1:30 Lunch
1:30-2:50 Causal learning (Gopnik, McClelland)

Causal and probabilistic learning and theory of mind
Alison Gopnik

When do human learners show indirect effects in contingency learning experiments?
James L. McClelland

2:50-3:20 Discussion (Moderator Josh Tenenbaum)
3:20-3:45 Coffee break, 2nd floor lounge, Corbett Hall
3:45-5:45 Language I (Aslin, Colunga, Saffran)

How a Bayesian ideal learner could acquire stationary and non-stationary structures
Richard Aslin

Modeling word learning in typical and atypical word learners
Eliana Colunga

Learning how words sound facilitates learning what words mean
Jenny Saffran

5:45-6:15 Discussion (Moderator Mark Johnson)
6:15-7:30 Dinner

Tuesday

7:00-8:45 Breakfast
8:45-10:05 Language II (Gerken, Newport, Lidz)

Is patience a virtue for a probabilistic learner?
LouAnn Gerken

Multiple statistics and maturational constraints on learning: Modeling both the complexity and simplicity of statistical language learning
Elissa L. Newport

10:05-10:30 Coffee break, 2nd floor lounge, Corbett Hall
10:30-11:10 Language II (continued)

The role of statistics in a selective theory of syntax acquisition
Jeff Lidz

11:10-11:40 Discussion (Moderator Terry Regier)
11:40-1:00 Lunch
1:00-2:20 Concepts, abstraction, and intuitive theories (Goodman, Tenenbaum)

Conceptual development as inductive programming
Noah Goodman

Learning to learn with hierarchical Bayesian models
Josh Tenenbaum

2:20-2:50 Discussion (Moderator Denis Mareschal)
2:50-3:20 Coffee Break, 2nd floor lounge, Corbett Hall
3:20-5:30 Poster spotlights (~10 minutes per person)
5:30-7:00 Dinner
7:00-9:00 Posters

Wednesday

7:00-8:45 Breakfast
8:45-10:05 Number, Actions, Objects (Kemp, Johnson, Baldwin, Lee)

Probabilistic principles of object perception
Charles Kemp

Brain and vision in infancy: Toward a neoconstructivist view of object and face perception
Scott P. Johnson

10:05-10:30 Coffee Break, 2nd floor lounge, Corbett Hall
10:30-11:50 Number, Actions, Objects (continued)

Action re-description
Dare Baldwin

A model of knower-level behavior in number-concept development
Michael D. Lee & Barbara W. Sarnecka

11:50-12:20 Discussion (Moderator Tom Griffiths)
12:20-1:20 Lunch
1:20 Group Photo; meet on the front steps of Corbett Hall
1:30 Guided tour of the Banff Centre

Rest of the afternoon: free time; dinner in town

Thursday

7:00-8:45 Breakfast
8:45-10:05 Language III (Chater, Griffiths, Perfors)

Language acquisition by simplicity, and the evolution of language
Nick Chater

Revealing inductive biases through iterated learning
Tom Griffiths

10:05-10:30 Coffee break, 2nd floor lounge, Corbett Hall

10:30-11:10 Language III (continued)

What's innate, and how much input is enough? Current views on poverty of stimulus arguments
Amy Perfors

11:10-11:40 Discussion (Moderator Rebecca Gomez)
11:40-1:00 Lunch
1:00-2:20 Social Cognition I (Schulz, Kushnir)

Where Piaget meets Vygotsky: How causal and social cognition interact to affect children's inferences and exploratory play
Laura Schulz

Statistical learning in a social world
Tamar Kushnir

2:20-2:50 Discussion (Moderator Fei Xu)
2:50-3:15 Coffee Break, 2nd floor lounge, Corbett Hall
3:15-4:35 Statistics Education (Lehrer, Ciancetta)

Improving statistics education: Children invent representations, measures, and models of variability
Richard Lehrer

Statistics students' reasoning when comparing distributions of data
Matthew Ciancetta

4:35-5:35 Discussion of educational implications (Moderator Tom Griffiths)
5:35-7:30 Dinner

Friday

7:00-8:30 Breakfast
8:30-9:50 Social Cognition II (Shafto, Shultz)

A Bayesian model of pedagogical reasoning
Pat Shafto

Simulation of infant development on false-belief tasks
Thomas Shultz

9:50-10:20 Discussion (Moderator Alison Gopnik)
10:20-10:40 Coffee Break, 2nd floor lounge, Corbett Hall
10:40-12:00 Philosophical implications (Danks, Strevens)

Computational models, cognitive development, and Hempel's problem
David Danks

Inferring probabilities from symmetries: The scientist as child
Michael Strevens

12:00-12:30 Discussion (Moderator Luca Bonatti)
12:30-1:30 Lunch

Checkout by 12 noon.

** 5-day workshops are welcome to use the BIRS facilities (2nd Floor Lounge, Max Bell Meeting Rooms, Reading Room) until 3 pm on Friday, although participants are still required to checkout of the guest rooms by 12 noon. **

Abstracts to follow (if desired) in alphabetical order by last name of speaker.

Probabilistic Models of Cognitive Development
May 25-29, 2009

ABSTRACTS

(in alphabetic order by speaker surname)

Speaker: Maxim Abelev, University of British Columbia (poster presentation)

Title: Learning constrained by learning: overhypothesis formation in young children

Abstract: The validity of inductive inference depends on the scope of generalization applied to particular observations. Is inductive generalization constrained by learners' history of prior inductive projections? Are property dimensions that prove useful as bases of generalization privileged in future learning? For induction to be self-constraining, learners must derive higher-order predictions or overhypotheses (e.g. object shape is predictive of object label) from first-order regularities (e.g. "ball" refers to spherical objects). We examine children's spontaneous overhypothesis formation by presenting them with multiple contingencies among object properties and testing their reasoning about novel contingencies involving the same property dimensions (e.g. a rock's colour and its chemical composition). To disentangle overhypothesis formation from children's prior beliefs, we vary the amount and the variability of evidence presented, as well as the intuitive plausibility of the target contingencies. Our findings suggest that the tendency to detect higher-order regularities is moderate overall and is stronger in those domains which children have weaker prior intuitions, and when the evidence presented is more variable. This is joint work with You-bin Park and Fei Xu.

Speaker: Dick Aslin, University of Rochester

Title: How a Bayesian ideal learner could acquire stationary and non-stationary structures

Abstract: A key question in any domain of pattern learning (language or vision) is how to acquire the underlying structural principles that allow for robust generalization. In a series of studies of adults in the domains of visual scenes and streams of speech we show that a Bayesian ideal learner provides a superior account of performance than a simple pair-wise associative learner. However, the Bayesian ideal learner operates in "batch mode" across an entire corpus of input, and human performance is both incremental and biased toward the early structural properties of the corpus. Alternative models will be described and implications for naive learners (infants and children) will be discussed.

Speaker: Dare Baldwin, University of Oregon

Title: Action redescription

Abstract: Action presents major challenges to processing: Motion tends to be rapid, complex in its changing relations to other things in the world, and evanescent. As observers, we frequently have access only to fragments of the actual motion stream; we must rapidly register what is relevant and carry out extensive analysis (such as categorization, integration, inference) on-the-fly. This analytic process must be carried out on many levels simultaneously: judgments about action require both

sensitivity to fine-grained spatial and temporal details, and the ability to meaningfully integrate lengthy sequences of action separated by long gaps in time. A common thread across all levels of analysis in action processing may be a joint requirement for detecting and redescribing structure. Recent research indicates that adults are skilled at capitalizing on structure inherent in action, and show a propensity to redescribe such structure to promote inferences. A structure detection/redescription framework for approaching investigation of action processing has the advantage of accounting for how adults process action while also offering a natural approach to illuminating the epigenesis of such skill. My working hypothesis is that structure-detection skills enable infants to become increasingly tuned to relevant information within the motion flow, while redescription skills drive increasingly sophisticated levels of encoding that promote inferences about functional commonalities, intentions and goals. I will report on several lines of research underscoring the promise of this approach.

Speaker: Luca L. Bonatti, University of Nantes, France

Title: Predicting the future without knowing the past: Infants' sensitivity to probabilities

Abstract: How do infants predict the next future event, when such prediction involves estimating its probability? Adult literature suggests that humans often fail this task because they are affected by heuristics and biases, or because they can reason about the frequency of classes of events but not about the probability of single events. I will present evidence suggesting that already at 12 months infants have an intuitive notion of probability that applies to single, never experienced, events, and that they use it to predict what will happen next. I will also show that infants can integrate the probabilistic and non-probabilistic cues afforded by a scene, modifying their initial expectations on the basis of its phisical properties. I will argue that infants' intuitive grasp of the probability of future events derives from their logical representation of future possibilities.

Speaker: Elizabeth B. Bonawitz, Massachusetts Institute of Technology (poster presentation)

Title: "The block makes it go!": Toddlers' ability to integrate prediction, action, and expectations about contact relations

Abstract: Some researchers have suggested that correlation information and information about action are bound in a single representation: "causal knowledge". If children have only observed correlation information, do they spontaneously try to generate the effect? Do they represent the relationship as potentially causal? I will present action and looking-time studies that suggest that even when toddlers (mean; 24 months) predict that one event will follow another, they neither initiate the first event to try to generate the second (as preschoolers, mean 47 months, do spontaneously), nor do they expect that the predictive relations will involve physical contact. Toddlers succeed at both of these inferences when the observation events are generated by an agent (Experiment 2) or described using causal language (Experiments 3 & 4). This suggests that information about agency and causal language play a role in helping children recognize the relationship between prediction, action, and contact causality. This is joint work with Darlene Ferranti, Ali Horowitz, Alison Gopnik, Andy Meltzoff, Jim Woodward, and Laura Schulz.

Speaker: Daphna Buchsbaum, University of California, Berkeley (poster presentation)

Title: Learning from statistical and causal cues to action structure

Abstract: Social reasoning depends on understanding the relationship between actions, goals and outcomes. Learners must take a continuous stream of observed behavior, divide it into distinct meaningful actions, and determine which of those actions lead to effects in the world. Prior work has

shown that both children and adults have surprisingly powerful statistical and causal learning abilities, which may play an important role in action understanding. Both children and adults can infer causal relationships from conditional probabilities, and probabilistic computational approaches, based on Bayesian statistics, have been shown to accurately model this ability across a wide variety of situations. Children are especially adept at learning the causal consequences of human actions, but how do they pick out those actions in the first place? Little is yet known about both the cues that are used to detect the boundaries between actions, and the computations that allow us to identify causal actions from within a longer sequence. Here we present a Bayesian analysis of action segmentation and causal inference, as well as a series of experiments with both adults and children, exploring how the ability to segment action and to infer its causal structure develops. This is joint work with Alison Gopnik, Tom Griffiths, and Dare Baldwin.

Speaker: Nick Chater, University College London

Title: Language acquisition by simplicity, and the evolution of language

Abstract: Children appear to learn language primarily from positive data---i.e., from observing what is said, and not what utterances are not acceptable. This raises a substantial problem of induction--how can such a rich and productive system be learned from apparently partial data. Drawing on joint work with Paul Vitanyi and Anne Hsu, I discuss how this inductive reasoning problem can be tackled by assuming that the learning follows a simplicity principle---choosing the 'model' of the language that provides the simplest encoding of the linguistic, or other, input. This approach allows an approximate quantification of how much linguistic input is required to learn particular linguistic regularities. I also discuss work, with Morten Christiansen and Florencia Reali, which considers the relationship between language acquisition and the evolution of language--arguing that language has been shaped, by processes of cultural evolution, to be especially easy to learn and process.

Speaker: Matthew Ciancetta, California State University, Chico

Title: Statistics students' reasoning when comparing distributions of data

Abstract: Results from a qualitative study designed to investigate university students' reasoning strategies as they were engaged in making informal statistical inferences about pairs of data sets will be presented and discussed. Participants included 275 university-level statistics students who completed a task-based web survey where they reasoned about data set comparisons set in various contexts. Six, in-depth, follow up interviews were analyzed to support and initially validate the findings from the surveys. An important result was the construction and refinement of an interpretive framework for reasoning about distributions of data. The framework was organized in a five-tiered lattice structure: Level 0 (Idiosyncratic); Level 1 (Local); Level 2 (Transitional); Level 3 (Initial Distributional); and Level 4 (Distributional). Across all groups of students there was a clear separation between those who could reason proportionally and those who could not, and that separation was correlated with students who appeared to view data distributionally and those who did not.

Speaker: Eliana Colunga, University of Colorado

Title: Modeling word learning in typical and atypical word learners

Abstract: In typical development, word learning goes from slow and laborious to seemingly rapid and effortless. Typically developing 3-year-olds are so skilled at learning noun categories that they seem to intuit the whole range of things in the category from hearing a single instance named, for example, consistently extending the name of a novel solid object to others of the same shape regardless of color,

material, texture. Both connectionist and Bayesian models have been proposed to account for this knowledge and how different sorts of representations, pre-existing constraints and experience may build it. In this talk I extend some of my previous modeling work to account for atypical word learning. Unlike typically developing children, late talkers (children below the 15th-20th percentile on productive vocabulary) do not consistently extend new names for objects by shape, and in fact some of them show a consistent preference for texture. Modeling these different patterns of performance requires different initial conditions and different experiences. I will argue that modeling these atypical patterns of learning and generalization provides critical insights as to the nature of the learner and the learning mechanisms.

Speaker: David Danks, Carnegie Mellon University

Title: Computational models, cognitive development, and Hempel's problem

Abstract: The classic formulation of Hempel's problem is to explain the extent (if any) that a white shoe confirms "All ravens are black." Introspectively, a non-black non-raven seems to provide no confirmation of the generalization; logically, such an observation does seem to provide some (small) confirmation. While Hempel's problem is oft-discussed in philosophical accounts of confirmation, I will argue that it (or variants) applies far more broadly. In particular, I will argue that almost all computational models of inductive inference are susceptible to a form of Hempel's problem. I will survey several natural responses, and argue that only one has the resources to solve the computational problem in a non-circular manner. I will also show how some recent work in cognitive development instantiates this response, and thereby has an answer to Hempel's problem. This response is not, however, a purely computational one. I thus conclude that we can solve Hempel's problem only by including non-computational constraints in our models.

Speaker: Stephanie Denison, University of British Columbia (poster presentation)

Title: Twelve to 14-month-old infants can predict single event probability with large sets

Abstract: We examined whether 12- to 14-month-old infants (n = 32) can reason about single event probability in large populations utilizing two different methodologies. Previous findings using a looking time paradigm revealed that infants could reason correctly about single event probability with small but not large set sizes (Teglas et al., 2007; Bonatti, 2008). Our first task used a looking time measure. We showed infants a large box filled with a 9:1 ratio of red to white ping-pong balls and removed a sample of either 1 red ball or 1 white ball. We found that 12-month-old infants looked approximately an equal amount of time at both outcomes, F (1,30) = 0.291, p = 0.59. This suggests that infants' expectations were not violated by either sample; therefore, using this measure infants do not seem to reason correctly about single-event probability. Our second task used a search task modeled after by Feigenson et. al (2004). The experimenter familiarized infants to two transparent jars, one containing a 4:1 ratio of pink to black lollipops, the other containing the opposite ratio. On the test trial, the experimenter removed one lollipop from each jar and placed them into separate opaque cups. She then covered the jars and infants crawled to one of the cups. Infants chose the cup associated with the jar containing the higher proportion of their preferred colour 78% of the time, significantly different from chance, t (31) = 2.12, p < .05. This suggests that infants can make predictions about single event probabilities involving large set sizes with a search task. Together these results suggest that infants can reason about single event probability if encouraged to reason predictively rather than post-dictively. This is joint work with Vanessa Waechtler and Fei Xu.

Speaker: Kathryn Dewar, University of British Columbia (poster presentation)

Title: Overhypothesis formation in 9-month-old infants

Abstract: This study examines a type of inductive learning: overhypothesis formation. Goodman (1955) used the following example to illustrate the idea. Suppose I show you some identical bags. From the first bag, I pull out some white marbles, from the second, some red marbles, and from the third, some green marbles. If I showed you a new bag, and pulled out one blue marble, what do you think the color of the next marble drawn would be? Your answer would probably be blue. The learner has formed a second-order generalization, or overhypothesis, that "Bagfuls of marbles are uniform in color" and it allows the learner to make predictions about this new bag with a new color of marbles. Here, we examine whether 9-month-olds are able to demonstrate this learning ability. Four boxes, with clear containers, were placed on a stage. An experimenter sampled four objects from boxes 1-3 and placed them into each box's container. Boxes contained multi-colored, same-shaped objects (Box1: cubes, Box2: spheres, etc.). In the experimental condition, from the test box, the experimenter sampled, one at a time, two identical objects of a novel shape (e.g., triangle) (expected outcome). In an unexpected outcome, the second object pulled from the test box was a previously seen shape (e.g., a cube). Sampling from boxes 1-3 support the overhypothesis that each box contains uniformly shaped objects. Sampling different-shaped objects (unexpected outcome), but not same-shaped objects (expected outcome), violates this overhypothesis. The control condition was identical, except on the unexpected outcome, the different-shaped object was sampled from its same-shaped box. Nine-month-olds looked longer to the unexpected outcome than the expected outcome in the experimental, but not the control, condition. This is joint work with Fei Xu.

Speaker: Michael C. Frank, Massachusetts Institute of Technology (poster presentation)

Title: Early word learning through communicative inference

Abstract: How do children learn their first words? While they are able to make use of distributional information about the co-occurrence of words and objects, even very young children also seem to take into account information about speakers' communicative intentions. Rather than being though of as purely statistical or purely social, children's early word learning is best modeled as a process of statistical inference about speakers' communicative intentions. Using an communicative inference framework allows our model to learn words accurately from natural corpus data, to explain a large range of developmental results, and to make novel developmental predictions. In addition, this framework offers insight into how the rich variety of non-linguistic information about speaker's intentions can be used in service of word learning. This is joint work with Noah Goodman and Josh Tenenbaum.

Speaker: LouAnn Gerken, University of Arizona

Title: Is patience a virtue for a probabilistic learner?

Abstract: One question that seems likely to differentiate among probabilistic models of development is how much total input is required for generalization. This question is complicated by the fact that very few input examples might allow a relevant generalization in some domains (e.g., where all strings demonstrate a particular property), whereas a larger subset of possible examples might be required in other domains (e.g., where the distribution of words in sentence frames is used to infer different grammatical categories). In my talk, I will discuss recent data suggesting that (1) infants are able to decide between alternative generalizations for a data set based on surprisingly little input; and (2) infants can determine the value of waiting for additional data in situations when the generalization

requires a larger data set.

Speaker: Vittorio Girotto, University IUAV of Venice, Italy

Title: Children's probabilistic reasoning about combinatorial possibilities

Abstract: In some recent studies, we showed that young children base their elementary probabilistic reasoning on an extensional estimation of possibilities (Girotto & Gonzalez, 2006; 2008; Teglas et al., 2007). In these studies, however, each possibility corresponded to an elementary event such as the withdrawing of a particular chip. Hence, the open question is whether children can draw correct probabilistic inferences when they have to mentally combine elementary events. In a new series of studies, we tried to answer this question by asking children of various ages and adults to make bets about combinations of elementary events. The results show that from the age of about 6 children can assess probabilities from combinations and that from the age of about 9 they attain the level of competence exhibited by adults. These results corroborate the extensional view of probabilistic reasoning: Naïve individuals judge the relative probability of compound events by comparing the relative numerosity of sets of compound events. However, they do not make a complete combinatorial analysis of all possibilities, and base their numerosity comparisons on some heuristic processes.

Speaker: Sharon Goldwater, University of Edinburgh (poster presentation)

Title: Online learning for Bayesian word segmentation

Abstract: Recently, Goldwater, Griffiths, and Johnson (2006; in press) proposed a Bayesian model of word segmentation in which the learner attempts to identify a lexicon and segmentation that can account for the observed input data (unsegmented phonemic transcripts) using a relatively small number of lexical items. Their model achieved good performance on a corpus of transcribed child-directed speech, and also showed high correlations with human performance in an experimental task (Frank et al., 2007). However, their learner used a batch learning procedure to identify the optimal solution under their model. In this work, we develop an online learning algorithm for the model of Goldwater et al. based on a standard method known as particle filtering. We explore the effects of varying the amount of memory the learner has available, and show that in the minimal-memory case, our algorithm shows striking similarities to the earlier online algorithm proposed by Brent (1999).

Speaker: Noah Goodman, Massachusetts Institute of Technology

Title: Conceptual development as inductive programming

Abstract: I will consider inductive inference of probabilistic programs as a metaphor for human conceptual development. I will use several case studies, including number and quantifier concepts, to illustrate this idea and how it extends more common notions of concept learning. These examples will suggest that humans rely on rich cross-situational evidence and on linguistic "placeholder structures" to learn highly structured concepts--and that the program induction framework explains this as coherent probabilistic inference.

Speaker: Alison Gopnik, University of California, Berkeley

Title: Causal and probabilistic learning and theory of mind

Abstract: In earlier work, we applied a causal Bayes net approach to instances of physical causality, like the "blicket detector". But children's theories of psychological causality are equally if not more

important. I will report several studies that investigate the role of causal learning in psychological understanding. In one series of studies we explore how children use causal learning to segment human action into meaningful units. Ina second series we explore the relation between imitative learning and causal learning. And in a final set of studies we explore how causal learning plays a role in the development of children's attributions of actions to either situations or traits.

Speaker: Tom Griffiths, University of California, Berkeley

Title: Revealing inductive biases through iterated learning

Abstract: Some of the most complex knowledge that people acquire, such as languages and concepts, can only be learned from other people, who themselves learned from other learners. The prevalence of this process of iterated learning raises a natural question: what are the consequences of iterated learning for the information being transmitted? I will present theoretical results indicating that, in the case where the learners are Bayesian agents, iterated learning converges to an equilibrium determined entirely by the inductive biases of those agents. These results constrain the use of iterated learning as an explanation for the properties of languages, and predict that the biases of human learners can be revealed by producing iterated learning in the laboratory. I will describe the results of some experiments testing this prediction, examining human inductive biases for functions and probability distributions related to language. This is joint work with Mike Kalish, Stephan Lewandowsky, and Florencia Reali.

Speaker: Scott P. Johnson, University of California, Los Angeles

Title: Brain and vision in infancy: Toward a neoconstructivist view of object and face perception

Abstract: Infants face an induction problem: From the myriad sensory inputs available, how do they identify what is important, and how do they build knowledge from these inputs? Debates on the solution to this problem are ancient and impassioned, yet relevant empirical work appeared only in the last century, when Piaget offered the first viable theory of infants' knowledge acquisition. The essence of Piaget's theory is constructivism: the building of complex knowledge from simpler perceptual and cognitive precursors with time and experience. More recently, the constructivist view was challenged by researchers dedicated to the idea of the "competent infant" and the possibility that some kinds of knowledge arise independently of experience. The competent-infant view has an intuitive appeal, but leaves unaddressed a vital challenge: to understand the mechanisms by which new knowledge arises. In this talk I will present a "neoconstructivist" approach to cognitive development that is informed by both these views and reflects recent advances in our knowledge of how innate learning mechanisms and visual experience combine in early development to yield increasingly complex representations of objects and faces.

Speaker: Charles Kemp, Carnegie Mellon University

Title: Probabilistic principles of object perception

Abstract: Before the age of 4 months, infants make inductive inferences about the motions of physical objects. Spelke and colleagues have suggested that the knowledge that guides these inferences can be characterized as a collection of categorical principles. We propose that probabilistic principles may provide a better way to characterize physical knowledge, and illustrate this proposal by describing a Bayesian model that incorporates probabilistic principles of rigidity and inertia. A probabilistic approach allows principles of object perception to be violated to a greater or lesser extent, and predicts that violations of different severity may elicit different behavioral responses. We discuss how this

prediction might be tested experimentally. This is joint work with Fei Xu.

Speaker: Tamar Kushnir, Cornell University

Title: Statistical learning in a social world

Abstract: Children learn about the world from other people. Thus, in order to model how children learn, it is critical to consider their developing social cognition. I will give several examples which demonstrate how social knowledge interacts with and guides young children's interpretation of statistical evidence generated by human actions. In study 1, preschoolers and toddlers interpret non-random (intentional) sampling as expressing a preference. In studies 2 and 3, preschoolers treat confounded evidence differently depending on who generates the actions – whether the actions are their own or someone else's (study 2) or whether the actor is knowledgeable or ignorant in a particular domain (study 3). All three studies establish that, in the absence of relevant social information, children make normative statistical inferences. Children's ability to evaluate statistical evidence in a social context may be an important contributor to their impressive learning abilities.

Speaker: Michael Lee, University of California, Irvine

Title: A model of knower-level behavior in number-concept development

Abstract: We develop and evaluate a model of behavior on the Give-N task, a commonly-used measure of young children's number knowledge. The model uses the knower-level theory of how children represent numbers. In doing the Give-N task, the model assumes children start out with a set of biases that make certain responses more likely a priori than others. These biases are inferred by the model, and are expressed as a base-rate. The model then assumes the base-rate is updated on each experimental trial in a way that depends critically on the interaction between the experimenter's request and the child's knower-level, which is also inferred by the model. Overall, the model provides an excellent fit to previously collected data from 82 children spanning the whole developmental range. We discuss ways in which our modeling approach can be extended to other developmental tasks, and can be used to help evaluate alternative theories of number representation against the knower-level theory.

Speaker: Rich Lehrer, Vanderbilt University

Title: Improving statistics education: Children invent representations, measures and models of variability

Abstract: The discipline of statistics develops models of chance processes that account for the variability in a wide variety of natural and social phenomena. Our research demonstrates that young students can engage in these practices in scientifically and mathematically rigorous ways. Students collect and organize real-world data in simple contexts of repeated measurement (e.g., every member of the class measures the arm-span of the teacher). These contexts recapitulate history in that some of the first efforts to model chance arose in contexts of measure, such as those by astronomers to account for the variability of observations of the same stellar distances. After measuring, students design representations of their collection of measurements. These invented representations reveal how the choices made by designers result in different shapes of the same data. Moreover, representing data invites consideration of how processes of measurement results in properties of the distribution of data (the data are usually symmetric with a large cluster of cases in the center). Students' invented representations often parallel those developed conventionally by the discipline, but students are now in a better position to understand conventions as reflecting (earlier) invention. Students go on to invent

measures of the "real" length of the object measured and the extent to which the measurements made by individual measurers tended to agree. These invented measures correspond to what the discipline terms statistics of center and variability. Assisted by new computer tools (Tinkerplots 2.0), students develop models of chance processes to account for the patterns in their data. Because data modeling is innovative and goes well beyond recall and use of formulas, the project designed new assessments to provide evidence of statistical reasoning. Four iterations of data modeling studies were conducted with fifth- and sixth-grade students in urban settings. The studies produced video of classroom interactions, collections of student artifacts, responses to assessment items, and interviews probing statistical reasoning. Each iteration was a re-design based on evidence obtained from findings in the previous iteration. Quasi-experimental comparisons showed that students engaged in modeling were more adept than peers at reasoning about probability, sample, and relations among different representations of the same data. Current work focuses on assessing students' statistical reasoning in ways that can productively inform classroom teaching and that spans a broader spectrum of reasoning (compared to all other standardized measures).

Speaker: Jeff Lidz, University of Maryland

Title: The role of statistics in a selective theory of syntax acquisition

Abstract: While research in the acquisition of syntax has largely focused on the necessity of abstract representations and the poverty of the stimulus with respect to these representations, very little research has asked how learners use the input to identify these representations. I present several experiments illustrating the role of statistical learning in a selective theory of syntax acquisition. I show (a) that infants can use statistical information to identify hierarchical phrase structure in an artificial grammar, (b) that the acquired representations allow for generalization to unobserved sentence structures, and (c) that statistical generalizations to be found in the input have consequences for morphosyntax that go beyond what can be inferred simply from the distributions. Hence, to the extent that learners use statistical information in learning syntax, they are doing so by comparing that information against the predictions of precise alternative syntactic representations.

Speaker: Christopher Lucas, University of California, Berkeley (poster presentation)

Title: A rational model of preference learning and choice prediction by children

Abstract: Young children demonstrate the ability to make inferences about the preferences of other agents based on their choices. However, there exists no overarching account of what children are doing when they learn about preferences or how they use that knowledge. We use a rational model of preference learning, drawing on ideas from economics and computer science, to explain the behavior of children in several recent experiments. Specifically, we show how a simple econometric model can be extended to capture two- to four-year-olds' use of statistical information in inferring preferences, and their generalization of these preferences. This is joint work with Tom Griffiths, Fei Xu, and Christine Fawcett.

Speaker: Lili Ma, University of British Columbia (poster presentation)

Title: Statistical inference in early social reasoning

Abstract: The present work examines whether infants and young children are able to use patterns of statistical evidence to make rational inferences in social reasoning, from two angles. In the first study, we found evidence that 9-month-olds attributed agency to a non-random sampling event that was psychologically less probable. When there was great regularity in the sample, infants expected to see a

human hand as the cause rather than a mechanical tool, as indicated by differences in their looking times. When the sampling was random, however, they did not have such an expectation. In the second study, we found evidence that by age 2 children were able to use sampling evidence to infer the subjective nature of preferences. After observing an adult sample only boring objects from a population that consisted of mostly interesting objects, 2-year-olds revised their prior belief that the adult would prefer the interesting objects as they themselves did and inferred that the adult actually preferred the boring objects. When the population consisted of only the boring objects, however, 2-year-olds continued to believe that the adult would prefer the interesting objects. Taken together, these two studies provide some of the first evidence for an early ability to make meaningful inferences from statistical evidence in social reasoning. This is joint work with Fei Xu.

Speaker: Denis Mareschal, Birkbeck College, University of London (poster presentation)

Title: The importance of the algorithmic level: The case of analogical completion

Abstract: Bayesian Models provide Computational Level descriptions of performance on a range of inductive inference and reasoning tasks. However, without a commitment to the algorithm that carries out the appropriate computations, they can say little about how the cognitive system breaks down under heavy cognitive loads or develops differently under atypical conditions. Such deviations are constrained by the algorithmic and implementational choices. We illustrate this point by describing a computational model of early analogical completion. The model suggests that much of analogical reasoning in young children and indeed much of everyday analogical reasoning reflects the priming in a distributed semantic memory systems and does not require and specialized analogical reasoning machinery. We argue that the model explains 6 key markers in the development of analogical reasoning

Speaker: James L. McClelland, Stanford University

Title: When do human learners show indirect effects in contingency learning experiments?

Abstract: In joint work, Daniel Sternberg and I have explored some factors influencing whether human learners show indirect effects (often called 'blocking' and 'screening') in contingency learning tasks. Inspired by Gopnik's experiments using the 'blicket detector', participants see one or two objects inside a frame. Later, a dot (the 'outcome') may or may not appear. Explicit prediction participants press a button to indicate their prediction on each trial, and only then is the outcome presented. For fast reaction participants, the dot does or does not appear exactly 350 msec after the objects, and participants must press a button within 250 msec of dot appearance. Both groups earn points for correct P or R responses, and loose points for incorrect responses. Half the participants in each group receive framing instructions similar to those used with children in Gopnik's studies, while the other participants are simply told to do their best to predict or react to the appearance of the dot. All participants learn the direct contingent relation between objects and outcomes. Indirect or inferred relations are not exhibited by participants in all groups.

Speaker: Elissa L. Newport, University of Rochester

Title: Multiple statistics and maturational constraints on learning: Modeling both the complexity and simplicity of statistical language learning

Abstract: Our recent empirical work on statistical language learning shows two important sets of findings relevant to modeling the learning process. First, studies of syntax learning in miniature languages - including verb-argument structure learning and the learning of grammatical categories - show that both adult and child learners are sensitive to a number of distinct statistics about the distributional

contexts they have heard for individual words and word categories, and that they coordinate these statistics in a sophisticated way to determine when to generalize and when to preserve lexical specificity. Second, comparisons between adult and child learners show that, under certain circumstances, children simplify and sharpen complex statistics, forming simpler and more regular rules as the outcome of learning. Ongoing work attempts to clarify the relationships between these two sets of findings, in order to model the child as learner and shaper of languages.

Speaker: Amy Perfors, University of Adelaide

Title: What's innate, and how much input is enough? Current views on poverty of stimulus arguments

Abstract: The issue of what type of innate biases we are born with (and, relatedly, how much input is necessary) is a deep and long-debated one, particularly with respect to questions of language acquisition. One common technique is to examine areas in which the input appears to be too impoverished to support the generalizations people make, and to use this observation to conclude that there must be a certain kind of innate knowledge permitting those generalizations. In this talk we revisit several standard "poverty of the stimulus" arguments in the light of modern computational techniques, which allow us to explore in more depth and specificity than before precisely how much it is actually possible to learn from different kinds of data. These techniques also have the advantage of forcing us as researchers to more precisely specify the learning problems facing young children and the assumptions we as theorists make about their capabilities -- as well as to determine exactly what the impact of those assumptions is. We conclude with a discussion of the broader implications of this work. This is joint work with Terry Regier and Josh Tenenbaum.

Speaker: Amy Perfors, University of Adelaide (poster presentation)

Title: Joint acquisition of word order and word reference

Abstract: Inferring the mappings between words and their referents is a difficult problem that all language learners face. Similarly, learning which word orders are permitted in one’s language is one of the first grammatical learning tasks these same learners must solve. We present a modeling framework which addresses simple versions of both of these problems by using the joint information in each to bootstrap the other. We discover that these two distinct learning tasks may be easier to solve jointly because of the way in which the inferences in one problem constrain the inferences in the other. This is joint work with Luke Maurits and Dan Navarro.

Speaker: Terry Regier, University of Chicago (poster presentation)

Title: Languages reflect near-optimal semantic categories

Abstract: Semantic categories in the world's languages appear to reflect both universal conceptual tendencies and linguistic convention. We argue that this mixed picture may be accounted for in terms of near-optimal partitions of semantic space - that is, partitions that near-maximize the dispersion of categories. We examine cross-language variation in the naming of categories from two semantic domains - color and space - and show that in both cases this notion of near-optimality accounts for universal tendencies while also accommodating some observed cross-language variation. This is joint work with Paul Kay, Naveen Khetarpal, and Asifa Majid.

Speaker: Jenny Saffran, University of Wisconsin-Madison

Title: Learning how words sound facilitates learning what words mean

Abstract: Despite a great deal of research effort focused on how infants learn about native language sounds, and an equally large research effort focused on how infants learn the meanings of words, these two areas are only beginning to make contact. In my talk, I will describe recent studies showing that what infants learn about the distributions of sounds influences subsequent word learning. Infants' early experiences with the statistics of sounds does not exist in a vacuum, but instead plays an important role in infants' subsequent acquisition of labels and their referents.

Speaker: Laura Schulz, Massachusetts Institute of Technology

Title: Where Piaget meets Vygotsky: How causal and social cognition interact to affect children's inferences and exploratory play

Abstract: I will discuss recent work looking at how prior knowledge and evidence affect children's exploratory behavior. I will then present three studies looking at how children's exploration of evidence is affected by social cues. The first study looks at how pedagogical information affects children's exploration of objects. The second study suggests that 16-month-olds exploration of objects is affected by their joint inferences about the projection of unobserved properties and the process by which the objects were sampled. The last study suggests that a context of contrastive beliefs (as compared to a context of contrastive possibilities) substantially improves children's ability to prove that one hypothesis is true and another false.

Speaker: Patrick Shafto, University of Louisville

Title: A Bayesian model of pedagogical reasoning

Abstract: Much of human learning takes place in pedagogical settings, settings where there is a person who chooses data for the purpose of teaching someone else a concept, and this has been argued to be an important di_erence between humans and other animals. I will present a computational model of pedagogical reasoning, which formalizes the problem and explains when and how teaching can facilitate learning. The pedagogical model treats teaching and learning as interrelated problems, generating predictions about which examples to choose in order to facilitate learning a concept, and what conclusions to draw from such examples. I will discuss data from experiments on adults showing that the model predicts intuitive teaching and learning about di_erent kinds of concepts, and quantifies how much more e_ective learning is in pedagogical situations. I will discuss the problem of identifying helpful teachers, implications of pedagogical demonstration for beliefs and exploration, and the hypothesis that pedagogy may form a “cultural ratchet”.

Speaker: Thomas Shultz, McGill University

Title: Simulation of infant development on false-belief tasks

Abstract: Onishi and Baillargeon (2005) used a violation-of-expectation paradigm to show that 15-month-olds succeed on a non-verbal version of the false-belief task, with looking time as a measure of surprise. A sibling-descendant cascade-correlation model covers the findings in this infant experiment and exhibits stages that had been observed in older children on verbal false-belief tasks: an omniscient theory of mind followed by a representational theory of mind, first for approach problems, and then for avoidance problems (Berthiaume et al., 2008, 2009). In contrast to current psychological explanations that emphasize inhibition and double inhibition (Leslie et al., 2005), network transitions

can be explained by learning probability distributions for where people search for missing objects. Contrasting this neural-network model with a Bayesian model (Goodman et al., 2006) highlights some interesting differences between these two approaches, particularly concerning issues of autonomous mental development and the building of cognitive structures.

Speaker: Michael Strevens, New York University

Title: Inferring probabilities from symmetries: The scientist as child

Abstract: Recent work by Xu and Garcia suggests that 8-month-olds are able to infer facts about physical probabilities from non-probabilistic facts, principally facts about physical symmetries. Philosophers have long recognized that adults make similar kinds of inferences. The philosophical focus has been on simple gambling setups, such as the drawings of balls from urns featured in Xu and Garcia's experiments. But some very important forms of inference in science, in statistical mechanics and evolutionary biology in particular, appear to have the same form. This paper explores the possibility that such inferences -- including inferences that provided the basis of a major scientific discovery -- are based in a perhaps innate capacity that manifests itself already in infants.

Speaker: Josh Tenenbaum, Massachusetts Institute of Technology

Title: Learning to learn with hierarchical Bayesian models

Abstract: In everyday learning and inference, people routinely draw successful generalizations from very limited evidence. Even young children can infer the meanings of words, the hidden properties of objects, or the existence of causal relations from just one or a few relevant observations -- far outstripping the capabilities of conventional learning machines. How do they do it? And how can we bring machines closer to these human-like learning abilities? I will argue that people's everyday inductive leaps can be understood as Bayesian inferences over hypothesis spaces generated by abstract knowledge representations -- what cognitive scientists have called "intuitive theories" or "schemas". A hierarchical Bayesian framework explains how abstract knowledge may itself be learned from experience, at the same as it functions to guide more specific generalizations from sparse data. I will discuss hierarchical Bayesian models for problems of "learning to learn" in the domains of categorization, word learning, and causal learning. A key challenge is to balance the need to hold strongly constrained inductive biases - necessary for rapid generalization - with the flexibility to adapt to the structure of new environments, learning new inductive biases for which we could not have been pre-programmed.

Speaker: Fei Xu, University of British Columbia

*Title***:** Rational statistical inference in infants

*Abstract***:** Are human infants rational learners? What are the statistical inference mechanisms available early in development, if any? I will report several series of experiments with infants between 6 and 12 months of age. Our results suggest that 1) during the second half of the first year, infants begin to be able to use the statistical properties of samples to make inferences about larger populations, and vice versa; 2) these inferential mechanisms can integrate domain-specific knowledge from both the physical and psychological realm; 3) different sampling conditions may license different kinds of inferences about objects and agents; and 4) infants' looking time patterns in these studies provide some preliminary evidence that they might be Bayesian learners. We suggest that the study of statistical inference mechanisms and the study of domain-specific knowledge must be integrated.